Energy Preserving Schemes for Nonlinear Hamiltonian Systems of Wave Equations. Application to the Vibrating Piano String.

Juliette Chabassier, Patrick Joly

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Energy Preserving Schemes for Nonlinear Hamiltonian Systems of Wave Equations. Application to the Vibrating Piano String.

Juliette Chabassier — Patrick Joly

N° 7168

Janvier 2010

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--7168--FR+ENG

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

centre de recherche
PARIS - ROCQUENCOURT

Energy Preserving Schemes for Nonlinear Hamiltonian Systems of Wave Equations. Application to the Vibrating Piano String.

Juliette Chabassier*†, Patrick Joly*‡

Thème : Modélisation et simulation numérique des ondes
Équipes-Projets POEMS

Rapport de recherche n° 7168 — Janvier 2010 — 70 pages

Abstract: The problem of the vibration of a string is well known in its linear form, describing the transversal motion of a string, nevertheless this description does not explain all the observations well enough. Nonlinear coupling between longitudinal and transversal modes seams to better model the piano string, as does for instance the "geometrically exact model" (GEM). This report introduces a general class of nonlinear systems, "nonlinear hamiltonian systems of wave equations", in which fits the GEM. Mathematical study of these systems is lead in a first part, showing central properties (energy preservation, existence and unicity of a global smooth solution, finite propagation velocity . . .). Space discretization is made in a classical way (variational formulation) and time discretization aims at numerical stability using an energy technique. A definition of "preserving schemes" is introduced, and we show that explicit schemes or partially implicit schemes which are preserving according to this definition cannot be built unless the model is linear. A general energy preserving second order accurate fully implicit scheme is built for any continuous system that fits the nonlinear hamiltonian systems of wave equations class.

Key-words: preserving schemes, energy, non linear systems of wave equations, piano string

* EPI POEMS
† juliette.chabassier@inria.fr
‡ patrick.joly@inria.fr

Centre de recherche INRIA Paris – Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Schémas numériques préservant une énergie pour les systèmes hamiltoniens non linéaires d'équations d'ondes. Application à la corde de piano.

Résumé : Le problème de vibration de corde est bien connu dans sa forme linéaire, où il décrit le mouvement transversal d'une corde. Cependant, cette description ne rend pas bien compte des observations. Un couplage non linéaire entre les modes transversal et longitudinal semble mieux adapté pour décrire la vibration d'une corde de piano, comme le fait par exemple le "modèle géométriquement exact" (MGE). Ce rapport introduit une classe générale de systèmes, les "systèmes hamiltoniens non linéaires d'équations d'ondes", dans laquelle entre le MGE. Dans une première partie, une étude mathématique de ces systèmes est menée, où l'on montre quelques propriétés essentielles (conservation d'une énergie, existence et unicité d'une solution gobale régulière, vitesse de propagation finie ...). La discrétisation en espace suit une méthode classique (formulation variationnelle) et la discrétisation en temps est menée de telle façon à atteindre la stabilité numérique grâce à une technique d'énergie. On introduit une définition de "schéma conservatif" et l'on montre que des schémas explicites ou partiellement implicites ne peuvent être conservatifs selon cette définition que si le modèle est linéaire. Un schéma numérique général, préservant l'énergie, précis à l'ordre deux, entièrement implicite, est donné pour n'importe quel système continu appartenant à la classe des systèmes hamiltoniens non linéaires d'équations d'ondes.

Mots-clés : schémas conservatifs, énergie, systèmes hamiltoniens non linéaires d'équations d'ondes, corde de piano

Contents

Introduction

Piano string vibration has been a recurrent subject of the scientific literature in acoustics. The problem of the vibration of a string is well known in its linear form, describing the transversal motion of a string, nevertheless this description does not explain all the observations well enough. Indeed, these observations illustrate a great complexity regarding the physical phenomena involved during the vibration. These phenomena are very numerous, as for example the hammer chock against the string, or the non linear coupling between transversal and longitudinal vibrations (longitudinal vibrations are in the same direction as the string). Several authors agree to say that the piano tone comes in particular from the longitudinal mode and its coupling with the transversal mode.

According to the linear theory, a vibrating string between two motionless points has a vibration spectrum composed with a fundamental frequency f_0 and harmonics frequencies equal to multiples of f_0. A term modeling the spectrum inharmonicity can be added to the linear model, describing the fact that harmonics are not worth exactly multiples of f_0 but a slightly higher value, the sharper the frequency, the bigger the inharmonicity. The harmonics are then called "partials". Inharmonicity can be quantified with an inharmonicity factor. Nakamura and Naganuma found in 1993 a second series of partials in piano sound spectra having one-fourth of inharmonicity compared to the main partial series ([31]). They attributed these to the horizontal polarization of the string. Giordano and Korty measured in 1996 that the longitudinal motion amplitude is a nonlinear function of the transversal amplitude ([13]), which confirms that the longitudinal vibration is generated by the transversal motion and not by the misalignment of the hammer, as some theories had put forward. Conklin measured again the partials found by Nakamura and Naganuma, and named them "phantom partials". He observed that the vibration relation between transverse and longitudinal modes greatly influences the tone quality ([8]), but also that the phantom partials are generated by a nonlinear coupling of these modes, noticing that the measured frequencies are sums or differences of the transversal model frequencies ([9]). Bank and Sujbert explained in 2005 that this result can be predicted by an approximated nonlinear model, coupling transversal and longitudinal modes ([2]). The model used in their paper is an approximation of the "geometrically exact" model introduced in [30]. The "geometrically exact" model comes from a geometric description of the string, a stress-strain relation and Newton's law. It is often exploited in the literature in an approximated form. The model and its approximations are presented in section (1).

All these phenomena, couplings, nonlinearities must be taken into account in numerical simulations aiming at reproducing a piano string sound. First numerical simulations have concerned a vibrating string in the transversal direction, not coupled with the longitudinal direction ([6]). The string was coupled with a nonlinear hammer model, which leads to interesting numerical results and a good fit with physical measurements. Nevertheless, the authors were not satisfied by the induced sound, attributing the lack of realism to the missing longitudinal mode. The sound of a piano string begins indeed with an audible high frequency chock, corresponding to the chock of the hammer on the string. It is transmitted very quickly to the rest of the structure through the longitudinal wave whose propagation speed is about ten times greater than the transversal one. Later simulations have consequently used coupled models. Several authors have dealt with coupled string models for the first transversal modes ([33, 26]) or using digital waveguides ([3], who made the approximation that the tension is uniform according to space). In [4] one can find a deep study of approximated nonlinear string models, an explanation of their origin regarding the "geometrically exact" model, a seek of good models ; and also linear numerical schemes for these models, with an energetic study of the schemes. The construction of numerical schemes was guided by the will to preserve an approximation of the energy preserved by the continuous system. This is a recurrent preoccupation when creating new schemes, for physical and stability reasons, but mostly particular or scalar equations are dealt with (ODE with polynomial nonlinearity [28, 24], ODE with general nonlinearity [17], EDP scalar equation with odd power [35], nonlinear scalar Klein Gordon equation [37, 11, 12, 5], more general EDP scalar equation [12], particular systems as Hénon-Heiles [29], nonlinear elasticity [15], ODE systems as a particle in a potential [16], or the 3 body problem [7]). In this study we will try to answer similar questions, in a general case of systems where no approximation is made regarding the model. We even consider any PDE system with a particular mathematical structure, with any number of unknown variables. This generality will allow our scheme to be applied to various fields, as elastodynamics where systems can have a very impressive size. Furthermore, it will cover all the models cited above.

The mathematical context of the geometrically exact model does not fit the usual hyperbolic theory of linear wave equations. We can introduce a general class of models which differ from each other in the expression of a function H, referred to as "potential energy". Properties of this potential energy lead to mathematical properties of the associated PDE system. We can indeed show, for instance, that the local hyperbolicity of the PDE system is equivalent to the local convexity of H. LiTaTsien theory ([27]) shows the existence of global classical solutions for locally hyperbolic quasilinear systems, which can then be applied when the potential energy is locally convex. Local convexity is achieved near the origin for H_{ex}, the potential energy of the geometrically exact model. Section (1) will describe some more model properties, specifying necessary conditions on the potential energy. Several authors have written the Taylor expansions of the geometrically exact potential energy around the origin, in order to work with simpler (polynomials) models : for instance, [1], [2] and [4] cited above. We will see how these expanded models are connected to the geometrically exact one, and wether or not they respect the properties displayed earlier.

The main contribution of our work, from the point of view of numerical analysis, is presented in section (2). We write a finite elements in space, finite difference in time, energy preserving scheme for a general class of models. Time discretization is tricky because of the nonlinearity of the models we consider. Schemes are built in order to preserve a discrete energy, consistent with the continuous one. We finally obtain a second order accurate, energy preserving scheme for any PDE system in our particular mathematical structure, with any number of unknown variables. At each time step, the nonlinear scheme is solved with an iterative Newton's algorithm. Section (3) presents some of the numerical difficulties and a few numerical results.

1 Non linear string models and their mathematical structure

Our aim in this study is to understand and model a piano string motion, thanks to the nonlinear coupling between different vibration directions. Motion equations lead to a particular form of PDE system, that we will call "hamiltonian system of wave equations". We will consider this form of PDE systems in very general point of view, by introducing a general class of systems differing from each other by the expression of a function H referred to as "potential energy". The "geometrically exact" model comes from the geometrical description of the string, a stress-strain relation and Newton's law. This model leads to a particular potential energy H_{ex} which satisfies numerous properties. Approximated models coming from a Taylor expansion of H_{ex}, for small values of its arguments, will be investigated, including the models used in [2] and [4]. We will lead a mathematical study of certain properties of this general class of models, under some assumptions on the potential energy H. First, let us wonder the physical origin of the model.

1.1 The geometrically exact model and its variants

1.1.1 Establishment of the geometrically exact model

We are interested in the string vibration, for instance a piano string. The problem has been formulated in its nonlinear version in [30], then used and modified by several authors. The geometrically exact model comes from a geometric description of the physics of the system and from mechanical fondamental laws. Figure 1, inspired by [2], presents the unknown variables of the problem. What follows is widely based on [36].

We will use the following notations:

- $x \in \Omega$ indicates the curvilinear coordinate along the string,
- Ω is a segment of $\mathbb{R}$ or $\mathbb{R}$ entirely (infinite string),
- $t > 0$ indicates time,
- $u(x,t)$ indicates the transversal component of the string motion (along $\mathbf{e}_y$),
- $v(x,t)$ indicates the longitudinal component of the string motion (along $\mathbf{e}_x$),
- E is the Young's modulus of the string,
- A is the section area,
- μ is the lineic mass of the string,
- T_0 is the rest tension.

We consider planar motion of a string[1], subjected to forces of tension resulting from its extension. The string is fixed at both ends, where its motion is zero. We assume that Young's modulus E, the string's section A, the lineic mass μ and the initial tension T_0 do not depend on x, although such a dependance could be possible in another model.

The position vector for a point marked by x is:

$$\mathbf{R}(x,t) = \big(v(x,t) + x\big)\mathbf{e_x} + u(x,t)\mathbf{e_y} = x\,\mathbf{e_x} + \mathbf{U}(x,t)$$

where $\mathbf{U}(x,t)$ is the vector of unknowns $\big(u(x,t), v(x,t)\big)$. For a given t, small δx, neglecting $O(|\delta x|^2)$, we have

$$\delta\mathbf{R}(x,\delta x,t) := \mathbf{R}(x,t) - \mathbf{R}(x+\delta x,t) \simeq \frac{\partial \mathbf{R}}{\partial x}\,\delta x \qquad \text{and} \qquad \Big|\delta\mathbf{R}(x,\delta x,t)\Big| \simeq \Big|\frac{\partial \mathbf{R}}{\partial x}\Big|\,\delta x$$

[1] A 3-D motion can be supposed, the resulting model will be presented at the end of this paragraph.

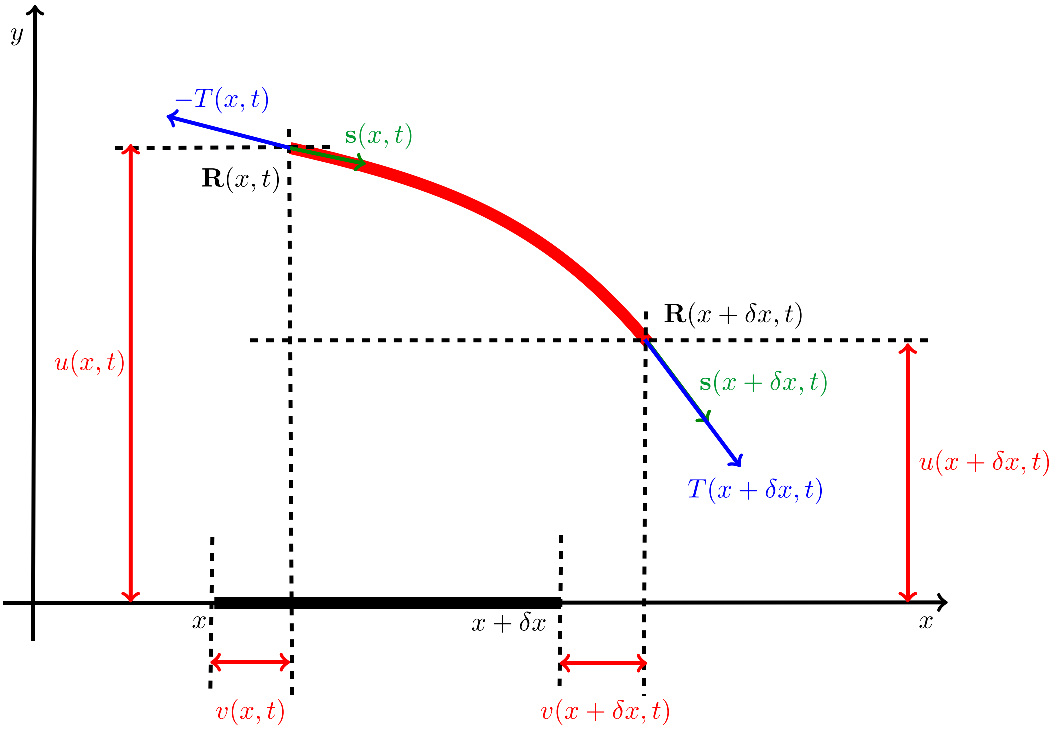


Figure 1: Transversal and longitudinal motions of a string.

with

$$\left|\frac{\partial \mathbf{R}}{\partial x}\right| = \sqrt{\left(\frac{\partial v}{\partial x}+1\right)^2+\left(\frac{\partial u}{\partial x}\right)^2}$$

Let us define $\mathbf{s}(x,t)$, the unit vector, tangent to the string at the point marked by x (see figure):

$$\mathbf{s}(x,t) = \frac{\frac{\partial \mathbf{R}}{\partial x}}{\left|\frac{\partial \mathbf{R}}{\partial x}\right|} = \frac{\left(\frac{\partial v}{\partial x}+1\right)\mathbf{e_x}+\left(\frac{\partial u}{\partial x}\right)\mathbf{e_y}}{\sqrt{\left(\frac{\partial v}{\partial x}+1\right)^2+\left(\frac{\partial u}{\partial x}\right)^2}}$$

We apply the second Newton's law to the string's segment $[x, x+\delta x]$. The only forces we take into account come from the tension of the string at the extreme points of the segment, directed along the unit vectors.

$$\frac{1}{\delta x}\mu\frac{d^2}{dt^2}\left(\int_x^{x+\delta x}\mathbf{R}(y,\delta x,t)\,dy\right) = \frac{1}{\delta x}\left[T(x+\delta x,t)\mathbf{s}(x+\delta x,t) - T(x,t)\mathbf{s}(x,t)\right]$$

hence, taking the limit when $\delta x \longrightarrow 0$,

$$\mu\frac{\partial^2 \mathbf{R}}{\partial t^2} = \frac{\partial}{\partial x}\left[T\mathbf{s}\right] \tag{1}$$

The physical stress-strain relation gives us an expression of the tension T, varying along the string, according to the deformation of the string, namely the relative extension $\delta a(x,t)$. The length of the element at rest is δx and becomes $\left|\mathbf{R}(x+\delta x,t)-\mathbf{R}(x,t)\right|$, hence the relative extension, after a Taylor expansion and neglecting $O(|\delta x|^2)$:

$$\delta a(x,\delta x,t) := \frac{\left|\mathbf{R}(x+\delta x,t)-\mathbf{R}(x,t)\right|-\delta x}{\delta x} \simeq \left|\frac{\partial \mathbf{R}}{\partial x}\right| - 1 \qquad \text{with } \frac{\partial \mathbf{R}}{\partial x} = \mathbf{e_x}+\frac{\partial \mathbf{U}}{\partial x}$$

In a general case, a stress strain law can be written:

$$T(x,t) = \phi\left(\delta a\,(x,t)\right) = \phi\left(\left|\mathbf{e_x} + \frac{\partial \mathbf{U}}{\partial x}(x,t)\right| - 1\right) \tag{2}$$

$\phi : \mathbb{R} \longrightarrow \mathbb{R}$ being a more or less pleasant function, normally growing, positive for any $x > 0$ and vanishing for $x = 0$. The system can then be written:

$$\mu \frac{\partial^2 \mathbf{U}}{\partial t^2} - \frac{\partial}{\partial x}\left[\phi\left(\left|\mathbf{e_x} + \frac{\partial \mathbf{U}}{\partial x}(x,t)\right| - 1\right)\ \frac{\mathbf{e_x} + \frac{\partial \mathbf{U}}{\partial x}}{\left|\mathbf{e_x} + \frac{\partial \mathbf{U}}{\partial x}\right|}\right] = 0$$

or, defining $d : \mathbb{R}^N \longrightarrow \mathbb{R}$ and $\mathbf{F} : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ as the following functions,

$$d(\mathbf{v}) = \left|\mathbf{e_x} + \mathbf{v}\right| - 1 \qquad \text{and} \qquad \mathbf{F}\big(\mathbf{v}\big) = \phi \circ d\big(\mathbf{v}\big)\ \frac{\mathbf{e_x} + \mathbf{v}}{\left|\mathbf{e_x} + \mathbf{v}\right|}$$

$$\boxed{\mu \frac{\partial^2 \mathbf{U}}{\partial t^2} - \frac{\partial}{\partial x}\mathbf{F}\Big(\frac{\partial \mathbf{U}}{\partial x}\Big) = 0} \tag{3}$$

This kind of systems can be written in a hamiltonian form if this function $\mathbf{F} : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is the gradient of another function. For our particular form of $\mathbf{F}$, this is the case. Indeed, we can show that

$$\nabla_{\mathbf{v}}\Big[\Phi \circ d\Big(\mathbf{v}\Big)\Big] = \mathbf{F}\Big(\mathbf{v}\Big) \tag{4}$$

Where $\Phi : \mathbb{R} \longrightarrow \mathbb{R}$ is an antiderivative of $\phi : \mathbb{R} \longrightarrow \mathbb{R}$. Indeed,

$$\nabla_{\mathbf{v}}\Big[\Phi \circ d\Big(\mathbf{v}\Big)\Big] = \nabla_{\mathbf{v}}\Big[d\big(\mathbf{v}\big)\Big]\ \Phi' \circ d\Big(\mathbf{v}\Big)$$

However, we can write that

$$d(\mathbf{v}) = \Big|\mathbf{e_x} + \mathbf{v}\Big| - 1 = \sqrt{\Big|\mathbf{e_x} + \mathbf{v}\Big|^2} - 1 = \sqrt{\ 1 + 2\,\mathbf{e_x} \cdot \mathbf{v} + \Big(\mathbf{v}\Big)^2} - 1$$

Hence its gradient:

$$\nabla_{\mathbf{v}}\Big[d\big(\mathbf{v}\big)\Big] = \frac{\mathbf{e_x} + \mathbf{v}}{\left|\mathbf{e_x} + \mathbf{v}\right|}$$

We conclude that if $\Phi' = \phi$, the formula (4) is correct. For any stress-strain law ϕ depending on the relative extension, we can see that the mechanical PDE system is always under the same form, which can lead to an hamiltonian form.

$$\mu \frac{\partial^2 \mathbf{U}}{\partial t^2} - \frac{\partial}{\partial x}\nabla\Big[\Phi \circ d\Big(\frac{\partial \mathbf{U}}{\partial x}\Big)\Big] = 0 \tag{5}$$

Setting $H = \Phi \circ d$, we obtain the very general form of system:

$$\boxed{\mu \frac{\partial^2 \mathbf{U}}{\partial t^2} - \frac{\partial}{\partial x}\nabla H\Big(\frac{\partial \mathbf{U}}{\partial x}\Big) = 0} \tag{6}$$

This form of systems will be referred to as “hamiltonian systems of wave equations”.

In the case where the stress-strain law is affine, it is called Hooke's law. The constant T_0 is called prestress of the string, and the law can be written:

$$\phi(\tau) = T_0 + EA\,\tau \qquad \text{or} \qquad T(x,t) = T_0 + EA\Big(\Big|\mathbf{e_x} + \frac{\partial \mathbf{U}}{\partial x}(x,t)\Big| - 1\Big) \tag{7}$$

Then, let us introduce $\Phi(\tau) = T_0\,\tau + EA\,\dfrac{\tau^2}{2}$, and the system can be written:

$$\mu\frac{\partial^2\mathbf{U}}{\partial t^2} - \frac{\partial}{\partial x}\Big[EA\big(\mathbf{e_x} + \frac{\partial\mathbf{U}}{\partial x}(x,t)\big) + (T_0 - EA)\frac{\mathbf{e_x} + \frac{\partial\mathbf{U}}{\partial x}(x,t)}{\big|\mathbf{e_x} + \frac{\partial\mathbf{U}}{\partial x}(x,t)\big|}\Big] = 0$$

Projecting the system on the axes and noticing that $\mathbf{e}_x$ does not depend on x, we obtain:

$$\left\{\begin{array}{l} \mu\,\dfrac{\partial^2 u}{\partial t^2} = \dfrac{\partial}{\partial x}\left[EA\,\dfrac{\partial u}{\partial x} - \;(EA - T_0)\,\dfrac{\frac{\partial u}{\partial x}}{\sqrt{\big(\frac{\partial u}{\partial x}\big)^2 + \big(1 + \frac{\partial v}{\partial x}\big)^2}}\right], \quad x \in \Omega, \quad t > 0, \\[4mm] \mu\,\dfrac{\partial^2 v}{\partial t^2} = \dfrac{\partial}{\partial x}\left[EA\,\dfrac{\partial v}{\partial x} - (EA - T_0)\,\dfrac{\big(1 + \frac{\partial v}{\partial x}\big)}{\sqrt{\big(\frac{\partial u}{\partial x}\big)^2 + \big(1 + \frac{\partial v}{\partial x}\big)^2}}\right], \quad x \in \Omega, \quad t > 0. \end{array}\right.$$

Finally, we have to add initial data (at $t = 0$) for u and v but also for $\dfrac{\partial u}{\partial t}$ and $\dfrac{\partial v}{\partial t}$; and boundary conditions at the string extremities. If the string is a segment $\Omega = [0, L]$, it is in a first time considered to be fixed at 0 and L, that is to say, $u(x = 0, t) = u(x = L, t) = 0$ and $v(x = 0, t) = v(x = L, t) = 0$.

With appropriate space and time scaling, we can write the following equivalent system, depending on a unique parameter $0 < \alpha < 1$ given by the formula $\alpha = \dfrac{EA - T_0}{EA}$.

$$\left\{\begin{array}{l} \dfrac{\partial^2 u}{\partial t^2} - \dfrac{\partial}{\partial x}\left[\dfrac{\partial u}{\partial x} - \alpha\,\dfrac{\frac{\partial u}{\partial x}}{\sqrt{\big(\frac{\partial u}{\partial x}\big)^2 + \big(1 + \frac{\partial v}{\partial x}\big)^2}}\right] = 0, \quad x \in \Omega, \quad t > 0, \\[4mm] \dfrac{\partial^2 v}{\partial t^2} - \dfrac{\partial}{\partial x}\left[\dfrac{\partial v}{\partial x} - \alpha\,\dfrac{1 + \frac{\partial v}{\partial x}}{\sqrt{\big(\frac{\partial u}{\partial x}\big)^2 + \big(1 + \frac{\partial v}{\partial x}\big)^2}}\right] = 0, \quad x \in \Omega, \quad t > 0. \end{array}\right. \tag{8}$$

Denoting

$$\forall\,(u_x, v_x) \in \mathbb{R}^2, \quad H_{ex}(u_x, v_x) = \frac{1}{2}\,u_x^2 + \frac{1}{2}\,v_x^2 - \alpha\,\Big[\sqrt{u_x^2 + (1 + v_x)^2} - (1 + v_x)\,\Big],$$

and

$$\mathbf{u} = (u, v),$$

we can write the string system in the same form of hamiltonian system of wave equations as (6):

$$\boxed{\frac{\partial^2\mathbf{u}}{\partial t^2} - \frac{\partial}{\partial x}\Big[\nabla H_{ex}\big(\frac{\partial\mathbf{u}}{\partial x}\big)\Big] = 0} \tag{9}$$

1.1.2 Geometrically exact model with three unknowns

We can generalize this geometrically exact model to the non planar motion of a string, considering two transversal displacements u_1 and u_2, and the longitudinal displacement v. The system of three equations can be derived in the same way as the previous system:

$$\left\{\begin{array}{lll}\dfrac{\partial^2 u_1}{\partial t^2} & - & \dfrac{\partial}{\partial x}\left[\dfrac{\partial u_1}{\partial x} - \alpha\,\dfrac{\frac{\partial u_1}{\partial x}}{\sqrt{\left(\frac{\partial u_1}{\partial x}\right)^2 + \left(\frac{\partial u_2}{\partial x}\right)^2 + \left(1+\frac{\partial v}{\partial x}\right)^2}}\right] = 0,\\[2ex] \dfrac{\partial^2 u_2}{\partial t^2} & - & \dfrac{\partial}{\partial x}\left[\dfrac{\partial u_2}{\partial x} - \alpha\,\dfrac{\frac{\partial u_2}{\partial x}}{\sqrt{\left(\frac{\partial u_1}{\partial x}\right)^2 + \left(\frac{\partial u_2}{\partial x}\right)^2 + \left(1+\frac{\partial v}{\partial x}\right)^2}}\right] = 0,\\[2ex] \dfrac{\partial^2 v}{\partial t^2} & - & \dfrac{\partial}{\partial x}\left[\dfrac{\partial v}{\partial x} - \alpha\,\dfrac{1+\frac{\partial v}{\partial x}}{\sqrt{\left(\frac{\partial u_1}{\partial x}\right)^2 + \left(\frac{\partial u_2}{\partial x}\right)^2 + \left(1+\frac{\partial v}{\partial x}\right)^2}}\right] = 0.\end{array}\right. \tag{10}$$

Denoting

$$H_{ex}(u_{1,x}, u_{2,x}, v_x) = \frac{1}{2}\,u_{1,x}^2 + \frac{1}{2}\,u_{2,x}^2 + \frac{1}{2}\,v_x^2 - \alpha\left[\sqrt{u_{1,x}^2 + u_{2,x}^2 + (1+v_x)^2} - (1+v_x)\right],$$

and

$$\mathbf{u} = (u_1, u_2, v),$$

we can also write the system (10) as (6).

1.1.3 Approximations of the geometrically exact model

Different Taylor expansions of H_{ex} with two unknowns lead to different approximated models. Let us notice that, near origin,

$$\left|\begin{array}{l}\sqrt{(1+v_x)^2 + u_x^2} = \sqrt{1+\eta(u_x,v_x)} = \\[1ex] \quad = 1 + \dfrac{1}{2}\,\eta(u_x,v_x) - \dfrac{1}{8}\,\eta(u_x,v_x)^2 + \dfrac{1}{16}\,\eta(u_x,v_x)^3 - \dfrac{5}{128}\,\eta(u_x,v_x)^4 + O(\eta(u_x,v_x)^5)\end{array}\right.$$

where

$$\eta(u_x, v_x) = 2\,v_x + u_x^2 + v_x^2$$

Developing and neglecting the terms in $O(\,|\mathbf{u}_x|^5\,)$, we have

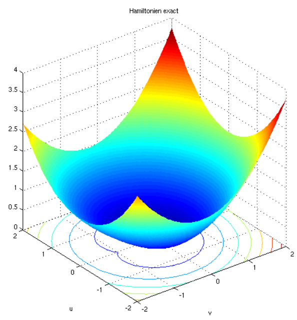
$$\left|\begin{array}{rcl}\sqrt{(1+v_x)^2 + u_x^2} & = & 1 + \dfrac{1}{2}\,\left(2v_x + u_x^2 + v_x^2\right) - \dfrac{1}{8}\,\left(2v_x + u_x^2 + v_x^2\right)^2 \\[1ex] & + & \dfrac{1}{16}\,\left(8v_x^3 + 12v_x^2(u_x^2+v_x^2)\right) - \dfrac{5}{128}\,(2v_x)^4 + O(\,|\mathbf{u}_x|^5\,)\end{array}\right.$$

that is to say

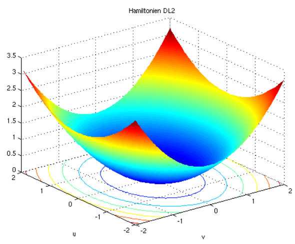
$$\sqrt{(1+v_x)^2 + u_x^2} - (1+v_x) = \frac{1}{2}\,u_x^2 - \frac{1}{2}\,u_x^2 v_x + \frac{1}{2}\,u_x^2 v_x^2 - \frac{1}{8}\,u_x^4 + O(\,|\mathbf{u}_x|^5\,)$$

Then H_{ex} verifies, near origin,

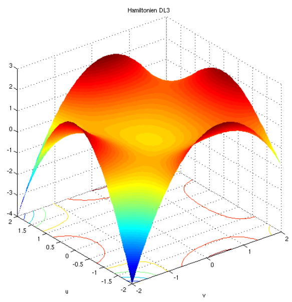
$$H_{ex}(u_x, v_x) = \frac{1}{2}\,u_x^2 + \frac{1}{2}\,v_x^2 - \alpha\Big[\frac{1}{2}\,u_x^2 - \frac{1}{2}\,u_x^2 v_x + \frac{1}{2}\,u_x^2 v_x^2 - \frac{1}{8}\,u_x^4\Big] + O(\,|\mathbf{u}_x|^5\,) \tag{11}$$



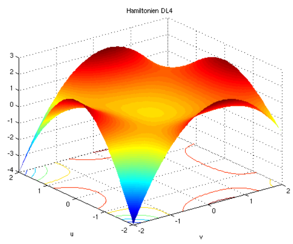
(a) Exact model



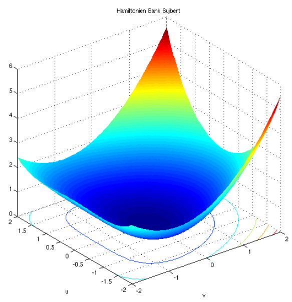
(b) Taylor order 2



(c) Taylor order 3



(d) Taylor order 4



(e) Bank-Sujbert model

Figure 2: Graphs of the different functions H

Linear model If we only consider the quadratic terms in (11), we make the approximation

$$\forall\,(u_x,v_x)\in\mathbb{R}^2,\quad H_{ex}(u_x,v_x)\simeq H_{DL2}(u_x,v_x)=\frac{1-\alpha}{2}\,u_x^2+\frac{1}{2}\,v_x^2,$$

and we write the classical linear model of two uncoupled wave equations:

$$\left\{\begin{aligned}&\frac{\partial^2 u}{\partial t^2}-(1-\alpha)\,\frac{\partial^2 u}{\partial x^2}=0,\quad x\in\Omega,\quad t>0,\\[2mm] &\frac{\partial^2 v}{\partial t^2}-\frac{\partial^2 v}{\partial x^2}=0,\quad x\in\Omega,\quad t>0.\end{aligned}\right.$$

From this system we can deduce an approximate propagation speed for each direction : 1 for the longitudinal direction, and $\sqrt{1-\alpha}$ for the transversal direction. Since $\alpha\simeq 1$, we acknowledge the well known result in mechanics saying that the longitudinal waves have a propagation speed much higher than the transversal waves. Figure 2(b) shows the graph of the second order of development.

Higher order models We can push further the Taylor expansion to superior orders:

$$H_{DL3}(u_x,v_x)=\frac{1-\alpha}{2}\,u_x^2+\frac{1}{2}\,v_x^2+\frac{\alpha}{2}\,u_x^2v_x$$

and

$$H_{DL4}(u_x,v_x)=\frac{1-\alpha}{2}\,u_x^2+\frac{1}{2}\,v_x^2+\alpha\Big[\frac{1}{2}\,u_x^2v_x-\frac{1}{2}\,u_x^2v_x^2+\frac{1}{8}\,u_x^4\,\Big]$$

Their graphs are plotted in figure 2(c) and figure 2(d).

An intermediary model The papers of Bank and Sujbert [2] and Bilbao [4] propose a less natural model, consisting in neglecting the quartic term $-\frac{\alpha}{2}\,u_x^2v_x^2$ in (11), which can be justified by a dimensional analysis in the case where the string is excited in the transversal direction (see appendix B). We obtain:

$$H_{BS}(u_x,v_x)=\frac{1-\alpha}{2}\,u_x^2+\frac{1}{2}\,v_x^2+\alpha\Big[\frac{1}{2}\,u_x^2v_x+\frac{1}{8}\,u_x^4\,\Big]$$

leading to the system

$$\left\{\begin{aligned}&\partial_t^2 u=\partial_x\Big[\,(1-\alpha)\,\partial_x u+\ \alpha\,(\partial_x u\,\partial_x v+\frac{1}{2}\,(\partial_x u)^3)\Big],\\ &\partial_t^2 v=\partial_x\Big[\,\partial_x v-\ \frac{\alpha}{2}(\partial_x u)^2\,\Big].\end{aligned}\right. \tag{12}$$

which will be referred to as “Bank and Sujbert” model during the following of this paper. The graph of this function can be seen in figure 2(e).

In this paragraph, we have built the geometrically exact model (8) and its approximations, governing a string motion with transversal-longitudinal coupling. We have shown that any stress-strain law, not necessarily affine as Hooke’s, leads to the same system structure (6) called “hamiltonian system of wave equations”. Even systems with higher number of unknowns as (10) can be written under this form. This structure will be considered in its total generality from now on. We will display some properties inherited by these systems under some hypothesis, and see if they apply to the geometrically exact model and its approximations.

1.2 General theoretical frame

This paragraph is devoted to the mathematical study of the hamiltonian systems of wave equations (6). We will see that the function H, which is a potential energy, totally determines the system. Geometrically exact model is associated with the energy H_{ex}, and the approximations coming from a Taylor expansion of the energy for small values of its arguments, at order n, are associated with the energies H_{DLn}. The mathematical properties we will find for the system will depend on hypotheses that must be satisfied by the potential energy. We will see if the geometrically exact model and its approximations satisfy the necessary hypotheses for all the properties.

1.2.1 General formulation

The geometrically exact model displayed in the last paragraph is an example of a general formulation of systems, called "hamiltonian systems of wave equations", which can be written, with Ω a segment of $\mathbb{R}$ or $\mathbb{R}$ entirely:

$$\boxed{\left\{\begin{array}{l}\text{Find } \mathbf{u} = (u_1, \cdots, u_N) : \Omega \times \mathbb{R}^+ \rightarrow \mathbb{R}^N, \\[2mm] \dfrac{\partial^2 \mathbf{u}}{\partial t^2} - \dfrac{\partial}{\partial x}\left[\nabla H\big(\dfrac{\partial \mathbf{u}}{\partial x}\big)\right] = 0, \quad x \in \Omega, \quad t > 0, \\[4mm] \mathbf{u}(x,0) = \mathbf{u}_0(x), \quad \dfrac{\partial \mathbf{u}}{\partial t}(x,0) = \mathbf{u}_1(x), \\[2mm] \mathbf{u}(x,t) = 0 \ \forall x \in \partial\Omega \end{array}\right.} \tag{13}$$

The function $H : \mathbb{R}^N \rightarrow \mathbb{R}$ stands for a potential energy.

Application to the geometrically exact string model. In the geometrically exact model, $N = 2$ and the function H, noted H_{ex} in the following, is given by:

$$\forall\ (u_x, v_x) \in \mathbb{R}^2, \quad H_{ex}(u_x, v_x) = \frac{1}{2}\ u_x^2 + \frac{1}{2}\ v_x^2 - \alpha\ \left[\sqrt{u_x^2 + (1 + v_x)^2} - (1 + v_x)\ \right], \tag{14}$$

Application to approximate models. In the approximate models, $N = 2$ and the functions H, coming from the Taylor expansions of H_{ex} at different orders, are given by:

$$\forall\ (u_x, v_x) \in \mathbb{R}^2, \quad H_{DL2}(u_x, v_x) = \frac{1-\alpha}{2}\ u_x^2 + \frac{1}{2}\ v_x^2 \tag{15}$$

$$\forall\ (u_x, v_x) \in \mathbb{R}^2, \quad H_{DL3}(u_x, v_x) = \frac{1-\alpha}{2}\ u_x^2 + \frac{1}{2}\ v_x^2 + \frac{\alpha}{2}\ u_x^2 v_x \tag{16}$$

$$\forall\ (u_x, v_x) \in \mathbb{R}^2, \quad H_{DL4}(u_x, v_x) = \frac{1-\alpha}{2}\ u_x^2 + \frac{1}{2}\ v_x^2 + \alpha\Big[\frac{1}{2}\ u_x^2 v_x - \frac{1}{2}\ u_x^2 v_x^2 + \frac{1}{8}\ u_x^4\ \Big] \tag{17}$$

$$\forall\ (u_x, v_x) \in \mathbb{R}^2, \quad H_{BS}(u_x, v_x) = \frac{1-\alpha}{2}\ u_x^2 + \frac{1}{2}\ v_x^2 + \alpha\Big[\frac{1}{2}\ u_x^2 v_x + \frac{1}{8}\ u_x^4\ \Big] \tag{18}$$

1.2.2 Energy preservation, H^1 stability

Hypothesis 1.1.1 *The function H is assumed positive.*

Remark 1.1 *The function H is only used through its gradient, hence any $H + c$ can fit for the equations, with $c \in \mathbb{R}$. Thus, its positivity is equivalent to being greater than a constant on $\mathbb{R}^2$. The following theorems are then still true if H is only greater than a constant.*

The PDE system (13) preserves an energy as stated in the following.

Theorem 1.1 *Any smooth enough solution* $\mathbf{u}$ *of (13) satisfies the energy identity:*

$$\frac{d}{dt}E(t) = 0, \quad \text{with } E(t) = \int_\Omega \left\{ \frac{1}{2} \left| \frac{\partial \mathbf{u}}{\partial t} \right|^2 + H(\frac{\partial \mathbf{u}}{\partial x}) \right\} dx. \tag{19}$$

PROOF. Let us multiply the equation with $\frac{\partial u}{\partial t}$ and integrate over x. We have:

$$\int_\Omega \left[\frac{\partial^2 \mathbf{u}}{\partial t^2} - \frac{\partial}{\partial x} \left[\nabla H(\frac{\partial \mathbf{u}}{\partial x}) \right] \right] \frac{\partial \mathbf{u}}{\partial t} = 0$$

Let us develop and integrate by parts the second part of the integral:

$$\int_\Omega \left[\frac{\partial^2 \mathbf{u}}{\partial t^2} . \frac{\partial \mathbf{u}}{\partial t} + \nabla H(\frac{\partial \mathbf{u}}{\partial x}) . \frac{\partial^2 \mathbf{u}}{\partial t \partial x} \right] = 0$$

Applying the rules of composed functions derivation, we have:

$$\int_\Omega \frac{1}{2} \frac{\partial}{\partial t} (|\frac{\partial \mathbf{u}}{\partial t}|^2) + \int_\Omega \frac{\partial}{\partial t} \left[H(\frac{\partial \mathbf{u}}{\partial x}) \right] = 0$$

Hence,

$$\frac{d}{dt} \int_\Omega \left\{ \frac{1}{2} \left| \frac{\partial \mathbf{u}}{\partial t} \right|^2 + H(\frac{\partial \mathbf{u}}{\partial x}) \right\} dx = 0$$

□

Remark 1.2 *If we consider the following equation, with a nonzero right member:*

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} - \frac{\partial}{\partial x} \left[\nabla H(\frac{\partial \mathbf{u}}{\partial x}) \right] = f, \quad x \in \Omega, \quad t > 0 \tag{20}$$

We just saw that if $f = 0$*, the system preserves an energy. A bound for the energy can be obtained when* f *is nonzero, namely:*

If H *is a positive function, the energy* $E(t)$ *verifies the following equation, for any* $t \geq 0$*:*

$$E(t) \leq \left[\sqrt{E(0)} + \frac{\sqrt{2}}{2} \int_0^t \left|\left| f(.,t) \right|\right|_{L^2_x} \right]^2$$

PROOF. *Let us multiply the equation by* $\frac{\partial u}{\partial t}$ *and integrate over* x*. We have:*

$$\frac{d}{dt} \int_\Omega \left\{ \frac{1}{2} \left| \frac{\partial \mathbf{u}}{\partial t} \right|^2 + H(\frac{\partial \mathbf{u}}{\partial x}) \right\} dx = \int_\Omega f(x,t) \frac{\partial \mathbf{u}}{\partial t}$$

Let $E(t) = \int_\Omega \left\{ \frac{1}{2} \left| \frac{\partial \mathbf{u}}{\partial t} \right|^2 + H(\frac{\partial \mathbf{u}}{\partial x}) \right\}$*. If* H *is positive, we can write the inequality:*

$$\left|\left| \frac{\partial \mathbf{u}}{\partial t} \right|\right|^2_{L^2_x} \leq 2\, E(t)$$

Then:

$$\begin{aligned} \frac{d}{dt}E(t) &= \int_\Omega f(x,t) \frac{\partial \mathbf{u}}{\partial t} \\ \frac{d}{dt}E(t) &\leq \left|\left| f(.,t) \right|\right|_{L^2_x} \left|\left| \frac{\partial \mathbf{u}}{\partial t} \right|\right|_{L^2_x} \\ \frac{d}{dt}E(t) &\leq \left|\left| f(.,t) \right|\right|_{L^2_x} \sqrt{2}\, \sqrt{E(t)} \end{aligned}$$

However, $\dfrac{dE}{dt} = 2\sqrt{E}\dfrac{d\sqrt{E}}{dt}$*, thus:*

$$\begin{aligned}
2\sqrt{E}(t)\frac{d\sqrt{E}}{dt}(t) &\leq \Big|\Big|f(.,t)\Big|\Big|_{L^2_x}\sqrt{2}\sqrt{E(t)} \\
\frac{d\sqrt{E}}{dt}(t) &\leq \frac{\sqrt{2}}{2}\Big|\Big|f(.,t)\Big|\Big|_{L^2_x} \\
\sqrt{E(t)} - \sqrt{E(0)} &\leq \frac{\sqrt{2}}{2}\int_0^t \Big|\Big|f(.,t)\Big|\Big|_{L^2_x} \\
E(t) &\leq \Big[\sqrt{E(0)} + \frac{\sqrt{2}}{2}\int_0^t \Big|\Big|f(.,t)\Big|\Big|_{L^2_x}\Big]^2
\end{aligned}$$

□

The energy preservation leads to an upper bound for H^1 norm of the solution, under a condition on H.

Theorem 1.2 *Let us assume that*

Hypothesis 1.2.1 *There exists $K > 0$ such that:*

$$\forall \mathbf{v} \in \mathbb{R}^N,\ H(\mathbf{v}) \geq K|\mathbf{v}|^2 \tag{21}$$

then, there exists $C > 0$ such that:

$$||\mathbf{u}(\cdot,t)||_{H^1} \leq C\,E(0), \qquad \forall t \geq 0$$

Proof.

$$\begin{aligned}
&\frac{d}{dt}\int_\Omega \left\{\frac{1}{2}\left|\frac{\partial \mathbf{u}}{\partial t}\right|^2 + H(\frac{\partial \mathbf{u}}{\partial x})\right\}\,dx = 0, \\
\Rightarrow\ &\int_\Omega \left\{\frac{1}{2}\left|\frac{\partial \mathbf{u}}{\partial t}\right|^2 + H(\frac{\partial \mathbf{u}}{\partial x})\right\}\,dx = E(0) \\
\Rightarrow\ &\int_\Omega H(\frac{\partial \mathbf{u}}{\partial x})\,dx = E(0) - \int_{\mathbb{R}} \frac{1}{2}\left|\frac{\partial \mathbf{u}}{\partial t}\right|^2\,dx, \\
\Rightarrow\ &K\int_\Omega \sum_{i=1}^N \Big(\frac{\partial u_i}{\partial x}\Big)^2 \leq \int_\Omega H(\frac{\partial \mathbf{u}}{\partial x})\,dx \leq E(0), \\
\Rightarrow\ &\Big|\Big|\frac{\partial u_i}{\partial x}\Big|\Big|^2_{L^2} \leq \frac{E(0)}{K}, \qquad \forall\, i \in [1,N].
\end{aligned}$$

Dirichlet boundary condition allows us to write Poincaré's inequality, and to conclude:

$$||u_i||_{H^1} \leq C\,E(0), \qquad \forall\, i \in [1,N].$$

□

Application to the geometrically exact string model. H^1 stability of the solution is then guaranteed if the function H is greater than a parabola. H_{ex} satisfies this condition, which leads to the H^1 norm stability of the solution of the geometrically exact model. Let us show that for any $(u_x, v_x) \in \mathbb{R}^2$,

$$(\mathcal{I}) \qquad H_{ex}(u_x,v_x) \geq \frac{1-\alpha}{2}(u_x^2 + v_x^2)$$

This inequality is equivalent to:

$$\begin{aligned}
& \frac{1}{2}u_x^2 + \frac{1}{2}v_x^2 - \alpha[\sqrt{u_x^2 + (1+v_x)^2} - (1+v_x)] \geq \frac{1-\alpha}{2}(u_x^2 + v_x^2) \\
\Leftrightarrow \quad & \frac{1-(1-\alpha)}{2}u_x^2 + \frac{1-(1-\alpha)}{2}v_x^2 \geq \alpha[\sqrt{u_x^2 + (1+v_x)^2} - (1+v_x)] \\
\Leftrightarrow \quad & \frac{\alpha}{2}u_x^2 + \frac{\alpha}{2}v_x^2 \geq \alpha[\sqrt{u_x^2 + (1+v_x)^2} - (1+v_x)] \\
\Leftrightarrow \quad & u_x^2 + v_x^2 + 2(1+v_x) \geq 2\sqrt{u_x^2 + (1+v_x)^2} \\
\Leftrightarrow \quad & u_x^2 + (1+v_x)^2 + 1 \geq 2\sqrt{u_x^2 + (1+v_x)^2} \\
\Leftrightarrow \quad & (\sqrt{u_x^2 + (1+v_x)^2} - 1)^2 \geq 0
\end{aligned}$$

This last inequality is always true, then $(\mathcal{I})$ is also true. Moreover, the constant $\dfrac{1-\alpha}{2}$ is optimal since the circle having $(0,-1)$ as a center and 1 as a radius makes the last equation vanish.

Application to approximate models. The approximate models coming from the 3^{rd} and 4^{th} order Taylor expansions of H_{ex} lead to potential energies H_{DL3} and H_{DL4} which are not always positive (see figures 2(c) and 2(d)). Consequently, hypothesis (1.2.1) cannot be satisfied, and these models do not guarantee stability if the energy is preserved.

The Bank-Sujbert model based on the potential energy H_{BS} satisfies hypothesis (1.2.1) with $K = \frac{1-\alpha}{2}$. Indeed, let us show that for any $(u_x, v_x) \in \mathbb{R}^2$,

$$(\mathcal{J}) \qquad H_{BS}(u_x, v_x) \geq \frac{1-\alpha}{2}(u_x^2 + v_x^2)$$

This inequality is equivalent to:

$$\begin{aligned}
& \frac{1}{2}u_x^2 + \frac{1}{2}v_x^2 - \alpha\left[\frac{1}{2}u_x^2 - \frac{1}{2}u_x^2 v_x - \frac{1}{8}u_x^4\right] \geq \frac{1-\alpha}{2}(u_x^2 + v_x^2) \\
\Leftrightarrow \quad & \frac{\alpha}{2}v_x^2 + \frac{\alpha}{2}u_x^2 v_x + \frac{\alpha}{8}u_x^4 \geq 0 \\
\Leftrightarrow \quad & v_x^2 + u_x^2 v_x + \frac{1}{4}u_x^4 \geq 0
\end{aligned}$$

This can be seen as a second order inequality in the unknown v_x.

$$\Delta = \left(u_x^2\right)^2 - 4 \times \left(\frac{u_x^4}{4}\right) = 0$$

Then the last inequality is true, and so is $(\mathcal{J})$. Moreover, the parabola described by v_x vanishes when $v_x = -\frac{1}{2}u_x^2$, which means that the constant $\dfrac{1-\alpha}{2}$ is optimal.

1.2.3 First order form and Hamiltonian structure

We want to show that the systems we study (hamiltonian systems of wave equations) can indeed be written under a hamiltonian form. Define $\mathcal{L} = -H$. This quantity is a lagrangian density, since it depends on x and t. The system can be written:

$$\partial_t^2 \mathbf{u} + \partial_x\big(\nabla\mathcal{L}(\partial_x \mathbf{u})\big) = 0$$

with $u : (\mathbb{R}^+ \times \Omega) \to \mathbb{R}^N$ and $\mathcal{L} : \mathbb{R}^N \to \mathbb{R}$. Setting $\mathbf{q} = \partial_x \mathbf{u}$ and $\mathbf{p} = \partial_t \mathbf{u}$, the system becomes:

$$\left\{\begin{array}{l} \partial_t \mathbf{p} = -\partial_x\big(\nabla\mathcal{L}(\mathbf{q})\big) \\[2mm] \partial_t \mathbf{q} = \partial_x \mathbf{p} \end{array}\right.$$

Introducing $\mathcal{H}(\mathbf{p},\mathbf{q}) = \frac{1}{2}\mathbf{p}^2 - \mathcal{L}(\mathbf{q})$, we have:

$$\begin{cases} \dfrac{\partial \mathcal{H}}{\partial \mathbf{q}} = -\nabla \mathcal{L}(\mathbf{q}) \\[2ex] \dfrac{\partial \mathcal{H}}{\partial \mathbf{p}} = \mathbf{p} \end{cases}$$

and the system can be written under a hamiltonian form:

$$\begin{cases} \partial_t \mathbf{p} = -\partial_x (\dfrac{\partial \mathcal{H}}{\partial \mathbf{q}}) \\[2ex] \partial_t \mathbf{q} = \partial_x (\dfrac{\partial \mathcal{H}}{\partial \mathbf{p}}) \end{cases}$$

Note that the physical hamiltonian, which can be written according to $\mathbf{u}$:

$$\mathcal{H}(\mathbf{u}) = \frac{1}{2}|\partial_t \mathbf{u}|^2 + H(\partial_x \mathbf{u})$$

is nothing but an energy density, as we have seen in previous paragraphs.

1.2.4 Hyperbolicity of the system

Definitions and mathematical results concerning hyperbolic systems of conservation laws can be found in [14]. We remind here some basic definitions.

Definition 1.1 Hyperbolic system
We consider the system of equations

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial}{\partial x}\mathbf{F}(\mathbf{u}) = 0 \tag{22}$$

where $\mathbf{F} : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}^n$ *and* $\mathbf{u} = (u_1, \cdots, u_n) : \Omega \times \mathbb{R}^+ \to \mathcal{D}$.

We define

$$\mathbf{A}(\mathbf{u}) = \left(\frac{\partial \mathbf{F}_i}{\partial u_j}\right)_{1 \leq i,j \leq n}$$

the Jacobian matrix of $\mathbf{F}$*. The system* (22) *is said to be hyperbolic if, for any* $\mathbf{u} \in \mathcal{D}$*, the matrix* $\mathbf{A}(\mathbf{u})$ *has* n *real eigenvalues* $\mu_1(\mathbf{u}) \leq \ldots \leq \mu_k(\mathbf{u}) \leq \ \ldots \leq \mu_n(\mathbf{u})$ *and* n *linearly independent corresponding eigenvectors* $\mathbf{r}_1(\mathbf{u}), \ldots, \mathbf{r}_k(\mathbf{u}), \ldots, \mathbf{r}_n(\mathbf{u})$*, i.e.*

$$\mathbf{A}(\mathbf{u})\, \mathbf{r}_k(\mathbf{u}) = \mu_k(\mathbf{u})\, \mathbf{r}_k(\mathbf{u}).$$

If, in addition, the eigenvalues $\mu_k(\mathbf{u})$ *are all distinct, the system* (22) *is called strictly hyperbolic.*

The system is said locally (strictly) hyperbolic if the appropriate properties are true not for $\mathbf{u} \in \mathcal{D}$ *but for* $\mathbf{u}$ *in a neighborhood of a certain point* $\mathbf{u}_0$.

Definition 1.2 Genuinely nonlinear field
The couple $\big(\mu_k(\mathbf{u}), \mathbf{r}_k(\mathbf{u})\big)$ *is said to be genuinely nonlinear or G.N.L. if*

$$\nabla \mu_k(\mathbf{u}) \cdot \mathbf{r}_k(\mathbf{u}) \neq 0 \ \forall \mathbf{u} \in \mathcal{D}$$

Definition 1.3 Linearly degenerate field
The couple $\big(\mu_k(\mathbf{u}), \mathbf{r}_k(\mathbf{u})\big)$ *is said to be linearly degenerate or L.D. if*

$$\nabla \mu_k(\mathbf{u}) \cdot \mathbf{r}_k(\mathbf{u}) = 0 \ \forall \mathbf{u} \in \mathcal{D}$$

Introducing

$$\mathbf{U} = (\frac{\partial \mathbf{u}}{\partial t}, \frac{\partial \mathbf{u}}{\partial x}) = (\mathbf{p}, \mathbf{q}) \in \mathbb{R}^{2N},$$

The system (13) can be written:

$$\begin{cases} \text{Find } \mathbf{U} : \Omega \times \mathbb{R}^+ \to \mathbb{R}^{2N}, \\ \dfrac{\partial \mathbf{U}}{\partial t} + \dfrac{\partial}{\partial x} F(\mathbf{U}) = 0, \quad x \in \Omega, \quad t > 0, \\ \mathbf{U}(x,0) = \mathbf{U}_0(x), \quad x \in \Omega. \end{cases} \tag{23}$$

where

$$\forall\, \mathbf{U} = (\mathbf{U}_t, \mathbf{U}_x) \in \mathbb{R}^N \times \mathbb{R}^N, \quad F(\mathbf{U}) = \begin{pmatrix} -\nabla H(\mathbf{U}_x) \\ -\mathbf{U}_t \end{pmatrix},$$

If we develop this expression for regular solutions, we obtain:

$$\begin{cases} \text{Find } \mathbf{U} : \Omega \times \mathbb{R}^+ \to \mathbb{R}^{2N}, \\ \dfrac{\partial \mathbf{U}}{\partial t} + A(\mathbf{U}) \dfrac{\partial \mathbf{U}}{\partial x} = 0, \quad x \in \Omega, \quad t > 0, \\ \mathbf{U}(x,0) = \mathbf{U}_0(x), \quad x \in \Omega. \end{cases}$$

where

$$\forall\, \mathbf{U} = (\mathbf{U}_t, \mathbf{U}_x) \in \mathbb{R}^N \times \mathbb{R}^N, \quad A(\mathbf{U}) = DF(\mathbf{U}) = \begin{pmatrix} 0 & -D^2 H(\mathbf{U}_x) \\ -I & 0 \end{pmatrix}.$$

where $D^2 H(\mathbf{U}_x)$ refers to the hessian matrix of H.

This form is usually referred to as "quasilinear system". This equation does not fit the "semilinear wave equations" context since the latters' nonlinearity should only depend on $\mathbf{u}$ and not its space derivatives as in (13). The theoretical context is very different.

Theorem 1.3 *Local hyperbolicity of the system (13) is equivalent to the local convexity of H.*

Proof. The eigenvalue problem

$$\text{Find } (\mathbf{Z}(\mathbf{U}) = (\mathbf{Z}_t(\mathbf{U}), \mathbf{Z}_x(\mathbf{U})), \mu(\mathbf{U})) \in \mathbb{C}^{2N} \times \mathbb{C}, \quad DF(\mathbf{U})\, \mathbf{Z}(\mathbf{U}) = \mu(\mathbf{U})\, \mathbf{Z}(\mathbf{U}),$$

is equivalent to

$$\text{Find } ((\mathbf{Z}_x(\mathbf{U}_x)), \mu(\mathbf{U})) \in \mathbb{C}^N \times \mathbb{C},$$

$$D^2 H(\mathbf{U}_x)\, \mathbf{Z}_x(\mathbf{U}) = \mu^2(\mathbf{U})\, \mathbf{Z}_x(\mathbf{U}) \quad \text{and} \quad \mathbf{Z}_t(\mathbf{U}) = -\mu(\mathbf{U}) \mathbf{Z}_x(\mathbf{U})$$

Thus, local hyperbolicity of the system (13) can only be achieved if H is locally convex.

□

Remark 1.3 *Local convexity of H leads to local hyperbolicity for the PDE system, which can allow us to use some hyperbolic existence and unicity results.*

Thus, the main problem admits $\mu(\mathbf{U})$ and $-\mu(\mathbf{U})$ as eigenvalues if and only if $\mu^2(\mathbf{U})$ is eigenvalue of $D^2H(\mathbf{U}_x)$. Since $D^2H(\mathbf{U}_x)$ only depends on $\mathbf{U}_x$, μ^2 and $\mathbf{Z}_x$ also only depend on $\mathbf{U}_x$. We try to solve:

$$D^2H(\mathbf{U}_x)\mathbf{v}(\mathbf{U}_x) = \lambda(\mathbf{U}_x)\mathbf{v}(\mathbf{U}_x)$$

Consequently, if $D^2H(\mathbf{U}_x)$ admits N eigenpairs $(\lambda_k, \mathbf{v}_k)_{1\le k\le N}$ then the first order formulation of the system will admit $2N$ eigenpairs :

$$\left(\left(\mu_k^+ := \sqrt{\lambda_k}, \mathbf{r}_k^+ := \begin{pmatrix} -\sqrt{\lambda_k}\mathbf{v}_k \\ \mathbf{v}_k \end{pmatrix}\right), \left(\mu_k^- := -\sqrt{\lambda_k}, \mathbf{r}_k^- := \begin{pmatrix} \sqrt{\lambda_k}\mathbf{v}_k \\ \mathbf{v}_k \end{pmatrix}\right) \right)_{1\le k\le N} \tag{24}$$

Application to the geometrically exact string model. H_{ex} is not a convex function since if we consider the aligned points:

$$\left| \begin{aligned} H_{ex}(-\alpha,-1) &= \frac{1-\alpha^2}{2} \\ H_{ex}(\ \ 0,-1) &= \frac{1}{2} \\ H_{ex}(+\alpha,-1) &= \frac{1-\alpha^2}{2} \end{aligned} \right.$$

the second image is greater that the other two, which contradicts the convexity of H_{ex}. Moreover, H_{ex} is not smooth (not even $\mathcal{C}^1$ near the point $u_x = 0, v_x = -1$). Anyway, we can consider the local $\mathcal{C}^\infty$ regularity and convexity of H_{ex} in a neighborhood of $(0,0)$, and the eigenvalues and eigenvectors of the system around this point.

Let us seek the eigenvalues and eigenvectors of $\nabla^2 H_{ex}(\mathbf{U}_x)$ for the nonlinear equations system when $N = 2$. We have:

$$\nabla H_{ex}(\mathbf{U}_x) = \nabla H_{ex}(u_x, v_x) = \begin{pmatrix} u_x \\ v_x \end{pmatrix} - \alpha\Big[\frac{1}{\sqrt{u_x^2 + (1+v_x)^2}}\begin{pmatrix} u_x \\ 1+v_x \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix}\Big]$$

Thus,

$$\nabla^2 H_{ex}(u_x, v_x) = \mathrm{Id}_{\mathbb{R}^2} - \frac{\alpha}{(\sqrt{u_x^2 + (1+v_x)^2})^3}\begin{pmatrix} (1+v_x)^2 & -u_x(1+v_x) \\ -u_x(1+v_x) & u_x^2 \end{pmatrix}$$

This matrix admits as eigenvalues $\lambda_1(u_x, v_x)$ and $\lambda_2(u_x, v_x)$:

$$\lambda_1(u_x, v_x) = 1 \qquad \text{and} \qquad \lambda_2(u_x, v_x) = 1 - \frac{\alpha}{\sqrt{u_x^2 + (1+v_x)^2}}$$

associated with the eigenvectors:

$$\mathbf{v_1}(u_x, v_x) = \begin{pmatrix} u_x \\ 1+v_x \end{pmatrix} \qquad \text{and} \qquad \mathbf{v_2}(u_x, v_x) = \begin{pmatrix} -(1+v_x) \\ u_x \end{pmatrix}$$

The first eigenvalue is then constant on the (u_x, v_x) plane, and the second eigenvalue is constant on circles centered on the point $(0,-1)$. Let us come back to the main problem. The eigenvalues coming from λ_1 are $\mu_1^+ = 1$ and $\mu_1^- = -1$, respectively associated with the eigenvectors $(-\mathbf{v_1}(u_x, v_x), \mathbf{v_1}(u_x, v_x))$ and $(\mathbf{v_1}(u_x, v_x), \mathbf{v_1}(u_x, v_x))$. The eigenvalues coming from λ_2 are

$$\mu_2^+ = \sqrt{1 - \frac{\alpha}{\sqrt{u_x^2 + (1+v_x)^2}}} \qquad \text{and} \qquad \mu_2^- = -\sqrt{1 - \frac{\alpha}{\sqrt{u_x^2 + (1+v_x)^2}}}$$

respectively associated with the eigenvectors

$$(-\mu_2^+\mathbf{v_2}(u_x, v_x), \mathbf{v_2}(u_x, v_x)) \quad \text{and} \quad (-\mu_2^-\mathbf{v_2}(u_x, v_x), \mathbf{v_2}(u_x, v_x)).$$

The system is then (strictly) hyperbolic as soon as μ_2^+ and μ_2^- are real (and distinct), i.e. as soon as $\sqrt{u_x^2 + (1+v_x)^2} \ge (>)\ \alpha$. This is true in particular near the point $(0,0)$, since $0 < \alpha < 1$.

Application to approximate Bank-Sujbert model. H_{BS} is not a convex function since if we consider the aligned points:

$$\left|\begin{array}{l} H_{BS}(-\sqrt{2\,\dfrac{1+\alpha}{\alpha}}, -\dfrac{2}{\alpha}) = \dfrac{2}{\alpha^2} - [1 + \dfrac{\alpha}{2} + \dfrac{1}{2\alpha}] \\ H_{BS}(\qquad 0 \qquad, -\dfrac{2}{\alpha}) = \dfrac{2}{\alpha^2} \\ H_{BS}(+\sqrt{2\,\dfrac{1+\alpha}{\alpha}}, -\dfrac{2}{\alpha}) = \dfrac{2}{\alpha^2} - [1 + \dfrac{\alpha}{2} + \dfrac{1}{2\alpha}] \end{array}\right.$$

the second image is greater that the other two for $0 < \alpha < 1$, which contradicts the convexity of H_{BS}. However, it is a polynomial function, then H_{BS} is $\mathcal{C}^\infty$ and we can express the eigenvalues and eigenvectors of $D^2 H_{BS}$.

$$\begin{cases} \lambda_1 = 1 - \alpha\Big[\dfrac{1}{2} - \dfrac{3}{4}u_x^2 - \dfrac{1}{2}v_x\Big] + \dfrac{\alpha}{2}\sqrt{\left(1 - v_x - \dfrac{3}{2}u_x^2\right)^2 + 4u_x^2} \\ \lambda_2 = 1 - \alpha\Big[\dfrac{1}{2} - \dfrac{3}{4}u_x^2 - \dfrac{1}{2}v_x\Big] - \dfrac{\alpha}{2}\sqrt{\left(1 - v_x - \dfrac{3}{2}u_x^2\right)^2 + 4u_x^2} \end{cases}$$

The eigenvectors are obtained as well:

$$\begin{cases} v_1 = \begin{pmatrix} -\dfrac{1}{2} + \dfrac{3}{4}u_x^2 + \dfrac{1}{2}v_x + \dfrac{1}{2}\sqrt{\left(1 - v_x - \dfrac{3}{2}u_x^2\right)^2 + 4u_x^2} \\ u_x \end{pmatrix} \\ v_2 = \begin{pmatrix} -\dfrac{1}{2} + \dfrac{3}{4}u_x^2 + \dfrac{1}{2}v_x - \dfrac{1}{2}\sqrt{\left(1 - v_x - \dfrac{3}{2}u_x^2\right)^2 + 4u_x^2} \\ u_x \end{pmatrix} \end{cases}$$

1.2.5 Existence of a global smooth solution

The theory of hyperbolic systems shows that in general, even if initial conditions are very smooth, there do not exist classical solutions beyond some finite time interval. Anyway, we will see in this paragraph how we can apply a global classical existence result found in the book of Li Ta-tsien [27], page 89, to the system under some conditions on H.

In [27], we consider the following Cauchy problem:

$$\begin{cases} \text{Find } \mathbf{U} : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^n, \\ \dfrac{\partial \mathbf{U}}{\partial t} + A(\mathbf{U})\dfrac{\partial \mathbf{U}}{\partial x} = 0, \\ \mathbf{U}(x, 0) = \mathbf{U}_0(x) \end{cases} \tag{25}$$

The hypothesis made on A are the following: we suppose that $A(\mathbf{U})$ is a $n \times n$ matrix, and that in a neighborhood of $\mathbf{U} = 0$, system (25) is strictly hyperbolic:

$$\lambda_1(0) < \lambda_2(0) < \cdots < \lambda_n(0).$$

Moreover we suppose that $\mathbf{U}_0$ is a $\mathcal{C}^1$ function of x with compact support:

$$\mathrm{Supp}(\mathbf{U}_0) \subseteq [\alpha_0, \beta_0]$$

Theorem 1.4 (Li Ta-tsien) *Suppose that $A(\mathbf{U})$ is $\mathcal{C}^2$ in a neighborhood of $\mathbf{U} = 0$. Suppose that furthermore that system (25) is linearly degenerate (see definition (1.3)). Then, let*

$$\theta = (\beta_0 - \alpha_0) \sup_{x \in \mathbb{R}} |\mathbf{U}_0^{'}(x)|,$$

there exists $\theta_0 > 0$ so small that for any given $\theta \in [0, \theta_0]$, Cauchy problem (25) admits a unique global solution $\mathbf{U} = \mathbf{U}(x,t) \in \mathcal{C}^1(\mathbb{R} \times \mathbb{R}^+)$.

Remark 1.4 *This result is shown in the book for systems with the property of "weak linear degeneracy", which is fulfilled when the system is linearly degenerated in the sense of Lax. Since the string model fits into linearly degenerated class of systems, we have simplified here the hypothesis of Li Ta-tsien's theorem. Corollary (4.1) of the book [27], page* 90*, refers to this simplification. However, the property of "weak linear degeneracy" is better suited to the problem since it is a necessary and sufficient condition. In the case where this property is not satisfied, for all size of initial data, the solution "blows up".*

In the case of hamiltonian systems of wave equations, the initial system is:

$$\begin{cases} \text{Find } \mathbf{u} : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^N, \\ \dfrac{\partial^2 \mathbf{u}}{\partial t^2} - \dfrac{\partial}{\partial x}\left[\nabla H(\dfrac{\partial \mathbf{u}}{\partial x})\right] = 0, \quad x \in \mathbb{R}, \quad t > 0, \\ \mathbf{u}(x,0) = \mathbf{u}_0(x), \quad \dfrac{\partial \mathbf{u}}{\partial t}(x,0) = \mathbf{u}_1(x). \end{cases} \tag{26}$$

We recall that it can be written in the form of system (25) with:

$$n = 2N, \qquad \mathbf{U} = (\frac{\partial \mathbf{u}}{\partial t}, \frac{\partial \mathbf{u}}{\partial x}) \in \mathbb{R}^n, \qquad \mathbf{U}_0 = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_0' \end{pmatrix}$$

and

$$\forall\, \mathbf{U} = (\mathbf{U}_t, \mathbf{U}_x) \in \mathbb{R}^N \times \mathbb{R}^N, \quad A(\mathbf{U}) = DF(\mathbf{U}) = \begin{pmatrix} 0 & -D^2 H(\mathbf{U}_x) \\ -I & 0 \end{pmatrix}.$$

In order to satisfy the hypothesis of theorem (1.4), conditions must be set on $\mathbf{u}_0$, $\mathbf{u}_1$, H. In order for $\mathbf{U}_0$ to be $\mathcal{C}^1$, $\mathbf{u}_1$ must be $\mathcal{C}^1$ and $\mathbf{u}_0$ must be $\mathcal{C}^2$. In order for $A(\mathbf{U})$ to be strictly hyperbolic in a neighborhood of $\mathbf{U} = 0$, H must be strictly convex in a neighborhood of $\mathbf{U}_x = 0$. In order for $\mathbf{U}_0$ to have a compact support, $\mathbf{u}_0$ and $\mathbf{u}_1$ must have a compact support. In order for $A(\mathbf{U})$ to be $\mathcal{C}^2$ in a neighborhood of $\mathbf{u}_x = 0$, $D^2 H$ must be $\mathcal{C}^2$, ie H must be at least $\mathcal{C}^4$ in a neighborhood of $\mathbf{U}_x = 0$.

Moreover, if we call $(\mu_k^{\pm}, \mathbf{r}_k^{\pm})_{1 \le k \le N}$ the eigenpairs of $A(\mathbf{U})$, and $(\lambda_k, \mathbf{v}_k)_{1 \le k \le N}$ the eigenpairs of $\mathcal{D}^2 H(\mathbf{U}_x)$ (linked by (24)), then

$$\nabla \mu_k^{\pm} \cdot \mathbf{r}_k^{\pm} = \pm \frac{\nabla \lambda_k}{2\sqrt{\lambda_k}} \cdot \mathbf{v}_k$$

Consequently, in order for the system to be linearly degenerated, the matrix $D^2 H(\mathbf{U}_x)$ must verify the property $\nabla \lambda_k(\mathbf{U}_x) \cdot \mathbf{v}_k(\mathbf{U}_x) = 0$ for any $1 \le\ k \le N$.

Theorem 1.5 *Suppose that H is $\mathcal{C}^4$ and strictly convex in a neighborhood of $\mathbf{u}_x = 0$. Assume also that $\mathbf{u}_0$ is $\mathcal{C}^2$, $\mathbf{u}_1$ is $\mathcal{C}^1$ and have both a compact support included in $[\alpha_0, \beta_0]$. Suppose furthermore that the hessian matrix $D^2 H$ is linearly degenerated. Then, let*

$$\theta = (\beta_0 - \alpha_0) \sup_{x \in \mathbb{R}} [|\mathbf{u}_0''(x)|, |\mathbf{u}_1'(x)|],$$

there exists $\theta_0 > 0$ so small that for any given $\theta \in [0, \theta_0]$, Cauchy problem (26) admits a unique global $\mathcal{C}^2$ solution $\mathbf{u} = \mathbf{u}(x,t)$ on $t \ge 0$.

Application to the geometrically exact string model. The function H_{ex} is strictly convex in a neighborhood of 0 as long as we stay far enough from the point $(0,-1)$. In this neighborhood, it is also $\mathcal{C}^\infty$. Let us determine if the eigenvalues of the main problem are linearly degenerate (LD). We remind (see page 19) that D^2H_{ex} has the following eigenvalues:

$$\lambda_1(u_x,v_x)=1 \qquad \text{and} \qquad \lambda_2(u_x,v_x)=1-\frac{\alpha}{\sqrt{u_x^2+(1+v_x)^2}}$$

associated with the eigenvectors:

$$\mathbf{v_1}(u_x,v_x)=\begin{pmatrix}u_x\\1+v_x\end{pmatrix} \qquad \text{and} \qquad \mathbf{v_2}(u_x,v_x)=\begin{pmatrix}-(1+v_x)\\u_x\end{pmatrix}$$

$A(\mathbf{u})$ admits $\mu_1^+=1$ and $\mu_1^-=-1$ as eigenvalues, respectively associated with the eigenvectors $(-\mathbf{v_1}(u_x,v_x),\mathbf{v_1}(u_x,v_x))$ and $(\mathbf{v_1}(u_x,v_x),\mathbf{v_1}(u_x,v_x))$, and

$$\mu_2^+=\sqrt{1-\frac{\alpha}{\sqrt{u_x^2+(1+v_x)^2}}} \qquad \text{and} \qquad \mu_2^-=-\sqrt{1-\frac{\alpha}{\sqrt{u_x^2+(1+v_x)^2}}}$$

respectively associated with the eigenvectors

$$(-\mu_2^+\mathbf{v_2}(u_x,v_x),\mathbf{v_2}(u_x,v_x)) \quad \text{and} \quad (-\mu_2^-\mathbf{v_2}(u_x,v_x),\mathbf{v_2}(u_x,v_x))$$

First of all, $\nabla_{\mathbf{U}}\mu_1^+=\nabla_{\mathbf{U}}\mu_1^-=(0,0,0,0)$. These eigenvalues μ_1^+ and μ_1^- are consequently LD. Now, let us consider μ_2^+ and μ_2^-. We notice first that they depend only on u_x and v_x, which yields:

$$\nabla_{\mathbf{U}}\mu_2^\pm=\begin{pmatrix}0\\0\\\partial_{u_x}\mu_2^\pm\\\partial_{v_x}\mu_2^\pm\end{pmatrix}=\begin{pmatrix}0\\\nabla_{\mathbf{U}_x}\mu_2^\pm\end{pmatrix}$$

However

$$\nabla_{\mathbf{U}_x}\mu_2^\pm=\pm\nabla_{\mathbf{U}_x}\left(\sqrt{\lambda_2}\right)=\pm\frac{\nabla_{\mathbf{U}_x}\lambda_2}{2\sqrt{\lambda_2}}$$

Where

$$\nabla_{\mathbf{U}_x}\lambda_2(u_x,v_x)=\frac{\alpha}{(\sqrt{u_x^2+(1+v_x)^2})^3}\begin{pmatrix}u_x\\1+v_x\end{pmatrix}$$

Then

$$\begin{aligned}\nabla_{\mathbf{U}_x}\mu_2^\pm\cdot\begin{pmatrix}-\mu_2^\pm\mathbf{v_2}(u_x,v_x)\\\mathbf{v_2}(u_x,v_x)\end{pmatrix} &= \pm\frac{\nabla_{\mathbf{U}_x}\lambda_2}{2\sqrt{\lambda_2}}\cdot\mathbf{v_2}(u_x,v_x)\\ &= \frac{1}{2\sqrt{\lambda_2}}\frac{\pm\alpha}{(\sqrt{u_x^2+(1+v_x)^2})^3}\begin{pmatrix}u_x\\1+v_x\end{pmatrix}\cdot\begin{pmatrix}-(1+v_x)\\u_x\end{pmatrix}\\ &= 0\end{aligned}$$

Thus,

$$\nabla\mu_2^+(u_x,v_x)\cdot v_2^+(u_x,v_x)=0,\ \forall u_x\neq 0 \text{ and } v_x\neq -1$$

and

$$\nabla\mu_2^-(u_x,v_x)\cdot v_2^-(u_x,v_x)=0,\ \forall u_x\neq 0 \text{ and } v_x\neq -1$$

These eigenvalues are then LD. Under assumptions on initial data, we can then apply theorem (1.5) and have a global $\mathcal{C}^2$ solution of the equation.

Remark 1.5 *The first eigenvalues μ_1^+ and μ_1^- suggest that we can find solutions of the equation associated to the eigenvalue 1 or -1 independently on the values taken by the solution, since $\mu_1^\pm$ do not depend on u_x and v_x. The propagation happens then with no deformation. Let c be in $\{-1, +1\}$. If we find $\mathbf{U}$ such that $\frac{\partial \mathbf{U}}{\partial x}$ is eigenvector associated with the eigenvalue c, then it is easy to see that:*

$$\frac{\partial \mathbf{U}}{\partial t} + DF(\mathbf{U})\frac{\partial \mathbf{U}}{\partial x} \Leftrightarrow \frac{\partial \mathbf{U}}{\partial t} + c\frac{\partial \mathbf{U}}{\partial x} = 0$$

Which is a linear transport equation for the vector $\begin{pmatrix} \mathbf{U}_t \\ \mathbf{U}_x \end{pmatrix}$ *and accepts the solution* $\begin{pmatrix} -c\mathbf{U}_0'(x-ct) \\ \mathbf{U}_0'(x-ct) \end{pmatrix}$.

Let us seek these solutions directly in the second order system, that is to say for the eigenvalue $c^2 = 1$: we are looking for $\mathbf{U}(x,t) = \big(u(x,t), v(x,t)\big)$ such that the propagation speed is 1. Setting $\mathbf{U}(x,t) = \big(u(x,t), v(x,t)\big) = \mathbf{U}_0(x-t) = \big(u_0(x-t), v_0(x-t)\big)$, we can write:

$$\begin{array}{lll} \partial_x \mathbf{U}(x,t) = \mathbf{U}_0'(x-t) & \textit{and} & \partial_t \mathbf{U}(x,t) = -\mathbf{U}_0''(x-t) \\ \partial_x^2 \mathbf{U}(x,t) = \mathbf{U}_0'(x-t) & \textit{and} & \partial_t^2 \mathbf{U}(x,t) = \mathbf{U}_0''(x-t) = \mathbf{U}_0''(x-t) \end{array}$$

We obtain the system:

$$\mathbf{U}_0''(x-t) - D^2 H_{ex}\big(\mathbf{U}_0'(x-t)\big)\mathbf{U}_0''(x-t) = 0$$

However, if $\mathbf{U}_0''$ is an eigenvector of $D^2H_{ex}(\mathbf{U}_0')$ associated with the eigenvalue 1, then

$$D^2 H_{ex}\big(\mathbf{U}_0'(x-t)\big)\mathbf{U}_0''(x-t) = \mathbf{U}_0''(x-t)$$

and

$$\mathbf{U}_0''(x-ct) - \mathbf{U}_0''(x-ct) = 0$$

Which confirms that $\mathbf{U}(x,t) = \big(u_0(x-t), v_0(x-t)\big)$ is a solution.

We must find $\mathbf{U}_0(z)$ such that:

$$\begin{pmatrix} u_0''(z) \\ v_0''(z) \end{pmatrix} = \begin{pmatrix} u_0'(z) \\ 1 + v_0'(z) \end{pmatrix}$$

Or equivalently,

$$\begin{pmatrix} u_0''(z) \\ v_0''(z) \end{pmatrix} = \gamma(z)\begin{pmatrix} u_0'(z) \\ 1 + v_0'(z) \end{pmatrix}$$

which is enough to know that $D^2H_{ex}\big(\mathbf{U}_0'(x-t)\big)\mathbf{U}_0''(x-t) = \mathbf{U}_0''(x-t)$. The two vectors are linked if their determinant vanishes:

$$\big(1 + v_0'(z)\big)u_0''(z) - v_0''(z)u_0'(z) = 0 \quad \Leftrightarrow \quad (1+v_0'(z))^2\left(\frac{u_0'(z)}{1+v_0'(z)}\right)' = 0 \qquad (27)$$

outside the points where $(1 + v_0'(z)) = 0$. Every functions (u_0, v_0) satisfying this last equality are then solutions propagating with a speed of 1. It is easy to display a particular, compact supported, pair of functions, by setting $u_0'(z)$ proportional to $\big(1 + v_0'(z)\big)$, for instance by taking $u_0(z)$ always equal to zero and $v_0(z)$ anything. Of course, a lot of other solitons (solutions propagating without deformation with the speed 1) can be found.

Application to approximate Bank-Sujbert model. The calculation of $\nabla\lambda.v$ leads to:

$$\begin{cases} \nabla\lambda_1.v_1 = \dfrac{\alpha}{8}\Big[9u_x^2 + 6\sqrt{\Delta} - 12\dfrac{1}{\sqrt{\Delta}} + 27\dfrac{1}{\sqrt{\Delta}}u_x^4 + 36\dfrac{1}{\sqrt{\Delta}}u_x^2 v_x + 9u_x^2 + 12\dfrac{1}{\sqrt{\Delta}}v_x^2 + 6v_x\Big] \\ \nabla\lambda_2.v_2 = \alpha\Big[\dfrac{9}{4}u_x^2 + \dfrac{3}{4}v_x - \dfrac{3}{8}\sqrt{\Delta} + \dfrac{3}{2}\dfrac{1}{\sqrt{\Delta}} - \dfrac{27}{8}\dfrac{1}{\sqrt{\Delta}}u_x^4 - \dfrac{9}{2}\dfrac{1}{\sqrt{\Delta}}u_x^2 v_x\Big] \end{cases}$$

where $\Delta = 4\Big[\big(1 - v_x - \frac{3}{2}u_x^2\big)^2 + 4u_x^2\Big]$.

This calculation shows that with this model, the "weak linear degeneracy" property is violated, since the system is genuinely nonlinear. This model does not preserve the linearly degenerated mathematical structure of the exact model, which is a major disadvantage of this approximation. Indeed, the blow up theorem[2] can be applied to this case, meaning that the $\mathcal{C}^2$ norm of the solution must blow up in a finite time depending on the size of the initial data.

Application to scalar nonlinear wave equations. This existence result is true for the geometrically exact model for $N \geqslant 2$ but is false when the equation is scalar ($N = 1$) for the nonlinear case. As the approximate Bank-Sujbert model mentioned earlier, the nonlinear scalar case contradicts the "weak linear degeneracy" property introduced by Li Ta-Tsien. Indeed, the blow up theorem[2] can be applied, meaning that the $\mathcal{C}^2$ norm of the solution must blow up in a finite time depending on the size of the initial data. This result has been mentioned before by John [21] and Kleinerman and Majda[25].

The nonlinear scalar wave equation often found in the literature is the following:

$$u_{tt} - \big(K(u_x)\big)_x = 0$$

It is possible to write this second order scalar equation as a first order system having two opposite eigenvalues $\lambda^\pm = \pm\sqrt{K'(u_x)}$ associated with the two eigenvectors

$$v^\pm = \begin{pmatrix} 1 \\ \mp\sqrt{K'(u_x)} \end{pmatrix}$$

Indeed the quantity $\nabla\lambda^\pm.v^\pm$ is:

$$\nabla\lambda^\pm.v^\pm = \pm\frac{K''(u_x)}{2\sqrt{K'(u_x)}}$$

The classical nonlinear string model is to set $K(v) = \dfrac{v}{\sqrt{1+v^2}}$, which means the model is not weakly linearly degenerated.

1.2.6 Finite propagation velocity

Theorem 1.6 *We assume that*

Hypothesis 1.6.1 *There exists a constant $C > 0$ such that*

$$\forall\, (\mathbf{u}_x) \in \mathbb{R}^N, \quad |\nabla H(\mathbf{u}_x)|^2 \leq 2\, C^2\, H(\mathbf{u}_x), \tag{28}$$

Then, the solution $\mathbf{u}$ *of the system* (13) *propagates with a velocity lower than* C.

Proof. Let us assume hypothesis 1.6.1. We can notice that this property induces that H is positive. We will use an energy technique. Let $V > 0$ be a speed, to be specified later, and $a \in \mathbb{R}$ such that the initial data have their support in $]-\infty, a[$. We have for all $t > 0$:

$$\int_{a+Vt}^{+\infty} \Big(\frac{\partial^2 \mathbf{u}}{\partial t^2} - \frac{\partial}{\partial x}\Big[\nabla H(\frac{\partial \mathbf{u}}{\partial x})\Big]\Big) \cdot \frac{\partial \mathbf{u}}{\partial t}\, dx = 0,$$

that is to say, after integration by parts,

$$\left|\begin{aligned} &\int_{a+Vt}^{+\infty} \frac{\partial}{\partial t}\Big(\frac{1}{2}\,|\frac{\partial \mathbf{u}}{\partial t}|^2\Big)\, dx \;+\; \int_{a+Vt}^{+\infty} \Big[\nabla H(\frac{\partial \mathbf{u}}{\partial x})\Big] \cdot \frac{\partial^2 \mathbf{u}}{\partial x \partial t}\, dx \\ &\qquad\qquad - \Big[\nabla H(\frac{\partial \mathbf{u}}{\partial x}) \cdot \frac{\partial \mathbf{u}}{\partial t}\Big](a+Vt, t) = 0, \end{aligned}\right.$$

[2]blow up theorem of Li Ta Tsien, see remark 1.4

which, introducing the energy density

$$\mathbf{e} = \frac{1}{2}\,|\frac{\partial \mathbf{u}}{\partial t}|^2 + H(\frac{\partial \mathbf{u}}{\partial x})$$

can be written

$$\int_{a+Vt}^{+\infty} \frac{\partial \mathbf{e}}{\partial t}\, dx - \Big[\ \nabla H(\frac{\partial \mathbf{u}}{\partial x}) \cdot \frac{\partial \mathbf{u}}{\partial t}\ \Big]\,(a+Vt,t) = 0$$

or, after derivation under the integral,

$$\frac{d}{dt}\int_{a+Vt}^{+\infty} \mathbf{e}\; dx + \Phi(a+Vt,t) = 0,$$

where

$$\Phi := V\;\mathbf{e} - \Big[\ \nabla H(\frac{\partial \mathbf{u}}{\partial x}) \cdot \frac{\partial \mathbf{u}}{\partial t}\ \Big].$$

Using the hypothesis on initial data, we have

$$\forall t>0, \qquad \int_{a+Vt}^{+\infty} \mathbf{e}(x,t)\; dx = -\int_0^t \ \Phi(a+Vs,s)\; ds.$$

Now we choose V big enough such that Φ is positive. This is possible since

$$|\,\nabla H(\frac{\partial \mathbf{u}}{\partial x}) \cdot \frac{\partial \mathbf{u}}{\partial t}\,| \ \leq\ \frac{V}{2}\,|\frac{\partial \mathbf{u}}{\partial t}|^2 + \frac{1}{2V}\,\Big|\nabla H(\frac{\partial \mathbf{u}}{\partial x})\Big|^2,$$

Consequently

$$\Phi \geq V H(\frac{\partial \mathbf{u}}{\partial x}) - \frac{1}{2V}\,\Big|\nabla H(\frac{\partial \mathbf{u}}{\partial x})\Big|^2 \geq \big(\,V - \frac{C^2}{V}\,\big)\; H(\frac{\partial \mathbf{u}}{\partial x})$$

thanks to the hypothesis 1.6.1. If we choose V=C, the function Φ is positive, and we have

$$\mathbf{e}(x,t) = 0 \quad \text{ for } \quad x > a+Vt, \quad t>0.$$

Since H is positive, we have

$$\mathbf{u}(x,t) = 0 \quad \text{ for } \quad x > a+Vt, \quad t>0.$$

This result show that the propagation speed of the solution of the Cauchy problem (13) is upper bounded by:

$$C = \left(\frac{1}{2}\ \sup_{(\mathbf{u}_x)} R(\mathbf{u}_x)\right)^{\frac{1}{2}}, \quad R(\mathbf{u}_x) := \frac{\Big|\nabla H(\mathbf{u}_x)\Big|^2}{H(\mathbf{u}_x)}. \tag{29}$$

$\square$

Application to the geometrically exact string model. Let us study the ratio $R(u_x,v_x)$ in the geometrically exact case, defined for $(u_x,v_x) \in D_R = \mathbb{R}^2 - \{(0,0);(0,-1)\}$, since the denominator vanishes in $(0,0)$ and the numerator is not defined in $(0,-1)$. If we find an upper bound for this ratio outside these points[3], then we can apply the theorem and conclude that the solution has a finite propagation velocity. We have:

$$\nabla H_{ex}(u_x,v_x) = (u_x,v_x)^t - \alpha\Big(\frac{u_x}{\sqrt{u_x^2+(1+v_x)^2}}\,,\ \frac{1+v_x}{\sqrt{u_x^2+(1+v_x)^2}} - 1\Big)^t$$

[3]see remark 1.6

Thus,

$$
\begin{aligned}
|\nabla H_{ex}(u_x,v_x)|^2 &= \left|u_x-\alpha\Big(\frac{u_x}{\sqrt{u_x^2+(1+v_x)^2}}\Big)\right|^2+\left|v_x-\alpha\Big(\frac{1+v_x}{\sqrt{u_x^2+(1+v_x)^2}}-1\Big)\right|^2\\
&= u_x^2+v_x^2-2\alpha\sqrt{u_x^2+(1+v_x)^2}+2+\frac{\alpha(1-\alpha)(1+v_x)}{\sqrt{u_x^2+(1+v_x)^2}}+2\alpha v_x+2\alpha^2
\end{aligned}
$$

Moreover,

$$
H_{ex}(u_x,v_x)=\frac{1}{2}u_x^2+\frac{1}{2}v_x^2-\alpha\Big(\sqrt{u_x^2+(1+v_x)^2}-(1+v_x)\Big)
$$

We want to show that : There exists $C>0$ such that, for any $(u_x,v_x)\in D_R$,

$$
R(u_x,v_x)\leq 2\,C^2,\qquad \text{ie,}\qquad |\nabla H_{ex}(u_x,v_x)|^2\leq 2\,C^2\;H_{ex}(u_x,v_x)
$$

However, this inequality can be written

$$
\begin{aligned}
&u_x^2+v_x^2-2\alpha\sqrt{u_x^2+(1+v_x)^2}+2\,\frac{\alpha(1-\alpha)(1+v_x)}{\sqrt{u_x^2+(1+v_x)^2}}+2\alpha v_x+2\alpha^2\\
&\qquad\leq 2\,C^2\Big[\frac{1}{2}u_x^2+\frac{1}{2}v_x^2-\alpha\Big(\sqrt{u_x^2+(1+v_x)^2}-(1+v_x)\Big)\Big]\\
\iff\quad &(1-C^2)(u_x^2+v_x^2)-2\alpha(1-C^2)\sqrt{u_x^2+(1+v_x)^2}+2\,\frac{\alpha(1-\alpha)(1+v_x)}{\sqrt{u_x^2+(1+v_x)^2}}\\
&+2\alpha(1-C^2)v+2\alpha(\alpha-C^2)\leq 0\\
\iff\quad &(1-C^2)\Big[(u_x^2+v_x^2)-2\alpha\sqrt{u_x^2+(1+v_x)^2}+2\alpha v\Big]+2\,\frac{\alpha(1-\alpha)(1+v_x)}{\sqrt{u_x^2+(1+v_x)^2}}\\
&+2\alpha(\alpha-C^2)\leq 0
\end{aligned}
$$

This last inequality is satisfied for $C=1$. Indeed, we have:

$$
\sqrt{u_x^2+(1+v_x)^2}\geq|1+v_x|\geq 1+v_x
$$

Multiplying by $2\alpha(\alpha-1)$ which is negative since $0<\alpha<1$, we obtain, for $(u_x,v_x)\in\mathbb{R}^2-(0,-1)$:

$$
2\alpha(\alpha-1)\sqrt{u_x^2+(1+v_x)^2}+2\alpha(1-\alpha)(1+v_x)\leq 0
$$

which is the inequality we want. Moreover, the supremum is reached in $(0,1)$:

$$
R(0,1)=\frac{|\nabla H_{ex}(0,1)|^2}{H_{ex}(0,1)}=2
$$

The propagation speed of the solution is then upper bounded by $C=1$ for the geometrically exact model.

Remark 1.6 *The inequality is true almost everywhere in* $\mathbb{R}^2$*, which ensures that the points* $(0,0)$ *and* $(0,-1)$ *are not "blowing up" points. We can show that* R *is not continuous in these two points, but stays bounded.*

Remark 1.7 *The second order Taylor expansion of* H_{ex}*, leading to a linear decoupled system, has shown (see page 12) that, in a first approximation, propagation velocities or transversal and longitudinal waves were* $\sqrt{1-\alpha}$ *and* 1*, which are both bounded by* 1*.*

Application to approximate Bank-Sujbert model. The potential energy associated with the approximate Bank-Sujbert model cannot satisfy hypothesis 1.6.1. Indeed,

$$|\nabla H_{BS}|^2(u_x, v_x) = \Big((1-\alpha)u_x + \alpha u_x v_x + \frac{\alpha}{2}u_x^3\Big)^2 + \Big(v_x + \frac{\alpha}{2}u_x^2\Big)^2$$

If we take this expression for $v_x = 0$ and see the resulting function of u_x, we obtain:

$$|\nabla H_{BS}|^2(u_x, v_x = 0) = \frac{\alpha^2}{4}u_x^6 + \left[\frac{\alpha^2}{4} + (1-\alpha)\alpha\right]u_x^4 + (1-\alpha)^2 u_x^2$$

which cannot be bounded by $H_{BS}(u_x, v_x = 0)$ which is an order 4 polynomial function of u_x.

1.2.7 Preservation of symmetries

The question asked in this paragraph is about symmetry preservation. Here symmetry means that components of the solution are odd or even. We wonder if, with symmetric initial data, the system will preserve these symmetries, and under what condition on the system and the function H. We will see that if a component is odd at the initial time, it stays odd, whatever is H ; and if a component is even at the initial time, it stays even if H is even according to this variable. Consider the system

$$\begin{cases} \text{Find } \mathbf{u} : \Omega \times \mathbb{R}^+ \to \mathbb{R}^N, & \\ \dfrac{\partial^2 \mathbf{u}}{\partial t^2} - \dfrac{\partial}{\partial x}\left[\nabla H(\dfrac{\partial \mathbf{u}}{\partial x})\right] = 0, & \forall\, x \in \Omega, \quad \forall\, t > 0 \\ \mathbf{u}(x,0) = \mathbf{u}_0(x), \quad \dfrac{\partial \mathbf{u}}{\partial t}(x,0) = \mathbf{u}_1(x), & \forall\, x \in \Omega \\ \mathbf{u}(x,t) = 0, & \forall\, x \in \partial\Omega \end{cases} \tag{30}$$

Theorem 1.7 *If initial data $u_{0,i}$ are odd for x for any $i \in [1, N]$, then the solution $\mathbf{u} = \mathbf{u}(x,t)$ of the system (30) preserves this property for x for any $t > 0$, whatever is H.*

If there exists a subset I_p of $\{1, \cdots, N\}$ such that for any $i \in I_p$, initial data $u_{0,i}$ are even for x ; for any $i \in I_p^c$, initial data $u_{0,i}$ are odd for x. If moreover H is an even function according to its variables x_i for any $i \in I_p$, then the solution $\mathbf{u} = \mathbf{u}(x,t)$ of the system (30) will present the same symmetries as the initial data : $u_i(x,t)$ will be even for x for any $i \in I_p$ and odd for x for any $i \in I_p^c$.

Proof. Let us assume first that the initial data are odd. We will see that the structure of the PDE system will preserve naturally this property, with no condition on H. Let us write the system.

$$\frac{\partial^2 \mathbf{u}}{\partial t^2}(x,t) - \frac{\partial}{\partial x}\left[\nabla H(\frac{\partial \mathbf{u}}{\partial x})\right](x,t) = 0$$

Or also

$$\begin{cases} \partial_t^2 \mathbf{u}(x,t) - \partial_x \mathbf{v}(x,t) = 0 & (31) \\ \mathbf{v}(x,t) = \nabla H(\partial_x \mathbf{u})(x,t) & (32) \end{cases}$$

Setting

$$\begin{cases} \tilde{\mathbf{u}}(x,t) := -\mathbf{u}(-x,t) \\ \tilde{\mathbf{v}}(x,t) := \nabla H(\partial_x \tilde{\mathbf{u}})(x,t) \end{cases}$$

We have, by the composed functions derivation:

$$\partial_x \tilde{\mathbf{u}}(x,t) = \partial_x \Big[- \mathbf{u}(-x,t) \Big] = \partial_x \mathbf{u}(-x,t) \tag{33}$$

hence:

$$\tilde{\mathbf{v}}(x,t) = \nabla H\big(\partial_x \mathbf{u}\big)(-x,t)$$

thus,

$$\partial_x \tilde{\mathbf{v}}(x,t) = -\partial_x \big[\nabla H\big(\partial_x \mathbf{u}\big)\big](-x,t) = -\partial_x \mathbf{v}(-x,t) \qquad \text{from (32)}$$

However, from the first line of the original system, taken in $(-x,t)$, we have:

$$\partial_t^2 \mathbf{u}(-x,t) = \partial_x \mathbf{v}(-x,t)$$

then:

$$\partial_x \tilde{\mathbf{v}}(x,t) = -\partial_t^2 \mathbf{u}(-x,t) = \partial_t^2 \tilde{\mathbf{u}}(x,t).$$

With the definition of $\tilde{\mathbf{v}}$,

$$\begin{cases} \partial_t^2 \tilde{\mathbf{u}}(x,t) - \partial_x \tilde{\mathbf{v}}(x,t) = 0, \\ \\ \tilde{\mathbf{v}}(x,t) = \nabla H(\partial_x \tilde{\mathbf{u}})(x,t). \end{cases} \tag{34}$$

That means that the couple $(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})$ is another solution of the system (31,32), and since the initial data are odd, $(\tilde{\mathbf{u}}, \tilde{\mathbf{v}})$ is solution of the same Cauchy problem as $(\mathbf{u}, \mathbf{v})$. Using the uniqueness of the solution, we conclude with no conditions on H that:

$$\mathbf{u}(x,t) = -\mathbf{u}(-x,t), \qquad \forall\, x \in \Omega \;,\; \forall\, t > 0.$$

Let us study now the case where some components of initial data of $\mathbf{u}$ are even (those whose indexes are in the set I_p). We can write, for $j \in [1,N]$:

$$\frac{\partial^2 u_j}{\partial t^2}(x,t) - \frac{\partial}{\partial x}\left[\partial_j H\Big(\frac{\partial \mathbf{u}}{\partial x}\Big)\right](x,t) = 0$$

Let us set

$$\tilde{u}_i(x,t) := u_i(-x,t) \quad \forall\, i \in I_p \qquad \text{and} \qquad \tilde{u}_i(x,t) := -u_i(-x,t) \quad \forall\, i \in I_p^c$$

First, we have:

$$\partial_t^2 \tilde{u}_j(x,t) = \partial_t^2 u_i(-x,t) \quad \forall\, j \in I_p \qquad \text{and} \qquad \partial_t^2 \tilde{u}_j(x,t) = -\partial_t^2 u_i(-x,t) \quad \forall\, j \in I_p^c$$

The difference with previous lines stands in the equality (33), which becomes:

$$\begin{array}{rcll} \partial_x \tilde{u}_i(x,t) &=& -\partial_x u_i(-x,t), & \forall\, i \in I_p, \\ \partial_x \tilde{u}_i(x,t) &=& \partial_x u_i(-x,t), & \forall\, i \in I_p^c. \end{array}$$

We can then write, for $j \in [1,N]$

$$\partial_x \Big[\partial_j H\Big(\big(\partial_x \tilde{u}_i\big)_{i\in I_p}, \big(\partial_x \tilde{u}_i\big)_{i\in I_p^c}\Big)\Big](x,t) = -\partial_x \Big[\partial_j H\Big(\big(-\partial_x u_i\big)_{i\in I_p}, \big(\partial_x u_i\big)_{i\in I_p^c}\Big)\Big](-x,t)$$

Now we want to find the equations of the system, and conclude thanks to unicity, that the solution is odd according to the directions that belong to I_p^c and even according to the directions that belong to I_p. For this, $\partial_j H$ must compensate the sign that appears in front of the second time derivative. A sufficient condition is that $\partial_j H$ is even along x_i for any $i \in I_p^c$, and odd along x_i for any $i \in I_p$.

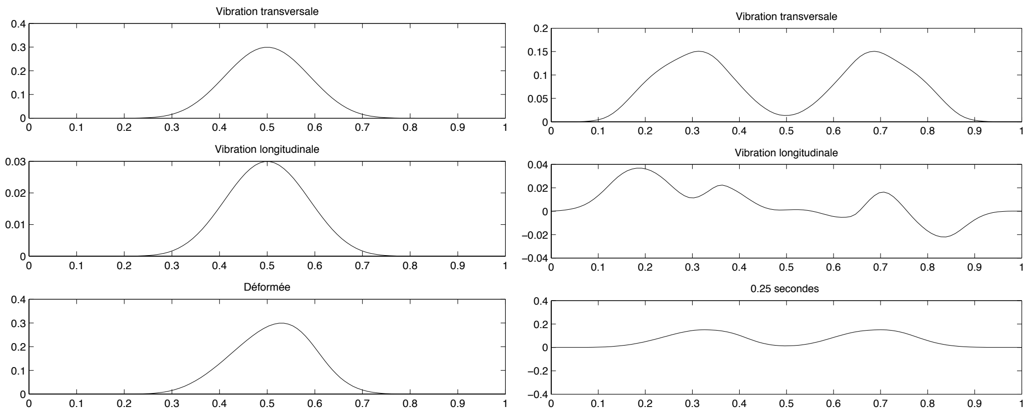
This condition is equivalent to H being even along the variables x_i for any $i \in I_p$. Then, we can write the original system in $(-x, t)$ and replace the expressions containing $\tilde{u}_i$:

$$\partial_t^2 \tilde{u}_j(x,t) = \partial_x \partial_j H\Big(\partial_x \tilde{\mathbf{u}}\Big)(x,t).$$

Then it appears that $\tilde{\mathbf{u}}$ is solution of the same system as $\mathbf{u}$, and given the initial conditions, it is solution of the same Cauchy problem. The uniqueness leads us to the conclusion that $\tilde{\mathbf{u}}(x,t) = \mathbf{u}(x,t)$, hence u_i is even for $i \in I_p$ and odd for $i \in I_p^c$.

To conclude, the function H must be even along its variables x_i for any $i \in I_p$ in order to preserve u_i even for any $i \in I_p$. On the opposite, the system's structure preserves u_i odd for any $i \in I_p^c$ without conditions on H.

□



(a) Initial Data

(b) After 0,25 seconds

Figure 3: Lost of the symmetry (E,E) with the geometrically exact model.

Application to the geometrically exact string model. The exact model takes as variables two directions, then it is possible to wonder about the preservation of four symmetries, denoting O the odd symmetry, and E the even : (E,E) , (E,O), (O,E) and (O,O). An odd symmetry will automatically be preserved by the structure of the PDE, as shown in the previous paragraph, but an even one will be preserved only if the function H_{ex} is even along the considered variable. Let us remind that:

$$H_{ex}(u_x, v_x) = \frac{1}{2}\, u_x^2 + \frac{1}{2}\, v_x^2 - \alpha\Big[\sqrt{(1+v_x)^2 + u_x^2} - (1+v_x)\Big]$$

Then, H_{ex} is even along u for any α, but is not even along v as soon as $\alpha > 0$. We can conclude that only the symmetries (E,O) and (O,O) will be preserved if $\alpha > 0$. On the opposite, if $\alpha = 0$, H_{ex} becomes even along all its variables, then all the symmetries are preserved with the linear model which appears.
Figure (3) shows a lost of the even symmetry along the longitudinal direction for $\alpha = 0,6$. The left subfigure (3(a)) shows the initial data (up : transversal direction, middle : longitudinal direction, down : string deformation) and the right subfigure (3(b)) shows the solution after $0,25$ seconds of numerical simulation in Matlab. One can notice the clear lost of even symmetry along longitudinal direction, whereas it is well preserved along transversal direction.

In this paragraph, we have studied the general class of systems (13) that we called "hamiltonian systems of wave equations", differing from each other in the expression of their potential energy H. The geometrically exact model fits into this class of systems, and so do its order 2, order 3, order 4 and Bank-Sujbert approximations.

We have seen that this type of systems preserved an energy (theorem 1.1), which could lead to H^1 stability of an eventual solution if the potential energy was greater than a parabola (theorem 1.2). The geometrically exact potential energy H_{ex} satisfies this property, as well as Bank-Sujbert potential energy H_{BS}. On the contrary, order 3 and order 4 approximations take negative values which contradicts the property.

We have written the general systems under a first order form, and shown that the local hyperbolicity of the system was equivalent to the local convexity of the potential energy (theorem 1.3).

As soon as H is $\mathcal{C}^4$ and locally strictly convex, existence and uniqueness of a global classical solution is possible (theorem 1.5) for small initial data if the hessian matrix of H is linearly degenerated (see definition 1.3). The geometrically exact model satisfies this LD property, hence it does not create singularities in the solution. On the opposite, Bank-Sujbert model is genuinely nonlinear (see definition 1.2), hence $\mathcal{C}^2$ singularities appear in the solution, in finite time, for any initial data.

Finite propagation velocity of the solution can be shown if the ratio between the gradient of the potential energy and itself is bounded (hypothesis 1.6.1). This bound is actually the bound on the propagation velocity. The geometrically exact model satisfies the hypothesis with a bound of 1, then its solution has a propagation velocity bounded by 1. On the contrary, Bank-Sujbert model presents no bound for the ratio.

Finally, if the initial data has odd components, the solution will be odd for these components. If the initial data has even components, the solution will be even for these components if H is even along these directions (see theorem 1.7).

2 Finite element energy preserving numerical schemes for nonlinear hamiltonian systems of wave equations

Each time one wishes to discretize in space and time an evolution problem whose solution satisfies the conservation of an energy, as it is the case of the systems (13) but more generally of many mechanical models, it is a natural idea to try to construct numerical schemes that preserve rigorously a discrete energy that is an equivalent of the continuous one. The first immediate interest is that one preserves after discretization a property of the exact solution, property that has an important physical meaning, particularly in mechanics. The second reason is that the stability of the scheme is generally ensured by the discrete energy, provided that this energy is positive. As we shall see immediately in the next paragraph, in the case of (13), the use of variational techniques (such as the finite element method) for the space semi-discretization ensures “by construction” the conservation of a positive semi-discrete energy. The difficulties really occur when the time discretization is concerned. This essential issue will be the object of the section 2.2.

2.1 Spatial semi discretization

2.1.1 Variational formulation

Let us consider the following system of partial differential equations:

$$\begin{cases} \partial_t^2 \mathbf{u} - \partial_x \left[\nabla H(\partial_x \mathbf{u})\right] = 0, \\[2ex] \mathbf{u}(x,t) = 0, \quad \forall\, t > 0, \forall\, x \in \partial\Omega. \end{cases} \tag{35}$$

Even though all what follows could probably be generalized to a more general context, we shall assume that the function H satisfies the coercivity property (21) and the additional assumption

$$\exists\, M > 0 \quad \text{such that} \quad |\nabla H(\mathbf{v})| \leq M\,(1 + |\mathbf{v}|), \quad \forall \mathbf{v} \in \mathbb{R}^N \tag{36}$$

In this case, according to the continuous energy identity, we expect that the solution u satisfies

$$u \in C^0(\mathbb{R}^+; H_0^1(\Omega)^N) \tag{37}$$

which implies, thanks to (36)

$$H(\partial_x u) \in L^\infty(\mathbb{R}^+; L^2(\Omega)^N) \tag{38}$$

In this framework, we can write a variational formulation in space of (13) in the space:

$$\mathcal{V} = \left(H_0^1(\Omega)\right)^N. \tag{39}$$

Let us take the inner product (in $\mathbb{R}^N$) of (13) by $\mathbf{v} \in \mathcal{V}$ and integrate over space the resulting equality. We get after integration by parts, since the boundary terms vanish,

$$\frac{d^2}{dt^2}\Big(\int_\Omega \mathbf{u}\cdot\mathbf{v}\Big) + \int_\Omega \nabla H(\partial_x \mathbf{u})\cdot\partial_x \mathbf{v} = 0, \quad \forall\, \mathbf{v} \in \mathcal{V} \tag{40}$$

which is the variational formulation of the problem.

2.1.2 Semi-discretization in space

We consider as usual $\{\,\mathcal{V}_h, h > 0\,\}$ a family of finite dimensional subspaces of $\mathcal{V}$, where h is an approximation parameter devoted to tend to 0. We assume the standard approximation property:

$$\forall\, \mathbf{v} \in \mathcal{V}, \quad \lim_{h\to 0}\ \inf_{\mathbf{v}_h\in\mathcal{V}_h} \|\mathbf{v} - \mathbf{v}_h\| = 0 \tag{41}$$

The most classical example is the approximation with conforming Lagrange finite element of degree $k \geq 1$, so-called P_k finite elements, on a family of meshes of Ω (in which case the approximation parameter is nothing but the stepsize of the mesh).

We consider the following semi-discrete problem: find $\mathbf{u}_h : \mathbb{R}^+ \mapsto \mathcal{V}_h$ such that

$$\frac{d^2}{dt^2}\left[\int_\Omega \mathbf{u}_h.\mathbf{v}_h\right] + \int_\Omega \nabla H(\partial_x \mathbf{u}_h).\partial_x \mathbf{v}_h = 0, \quad \forall\, \mathbf{v}_h \in \mathcal{V}_h\,. \tag{42}$$

We can write an algebraic formulation of (42) after having introduced the vector $\mathbf{U}_h \in \mathbb{R}^{N_h}$ (resp. $\mathbf{V}_h \in \mathbb{R}^{N_h}$) of the components of $\mathbf{u}_h$ (resp. $\mathbf{v}_h$) in an appropriate basis of $\mathcal{V}_h$. We first introduce the linear operator in $\mathbb{R}^{N_h}$, M_h and D_h^* defined by:

$$\Big(M_h \mathbf{U}_h, \mathbf{V}_h\Big)_h = \int_\Omega \mathbf{u}_h.\mathbf{v}_h, \quad \forall\, \mathbf{v}_h \in \mathcal{V}_h\,. \tag{43}$$

By analogy with the formula

$$\int_\Omega -\partial_x \nabla H(\partial_x \mathbf{u})\cdot \mathbf{v} = \int_\Omega \nabla H(\partial_x \mathbf{u})\cdot \partial_x \mathbf{v}$$

we introduce the nonlinear function in $\mathbb{R}^{N_h}$ (the complicated notation is chosen for convenience to emphasize the analogy with the continuous case - note that $-\partial_x$ is the formal adjoint of ∂_x)

$$\mathbf{D}_h^*\big(\nabla H(\mathbf{D}_h)\big) : \mathbf{U}_h \mapsto \mathbf{D}_h^*\big(\,\nabla H(\mathbf{D}_h \mathbf{U}_h)\,\big), \tag{44}$$

defined by

$$\Big(\mathbf{D}_h^*\big(\,\nabla H(\mathbf{D}_h \mathbf{U}_h)\,\big), \mathbf{V}_h\,\Big)_h = \int_\Omega \nabla H(\partial_x \mathbf{u}_h)\cdot \partial_x \mathbf{v}_h, \quad \forall\, \mathbf{v}_h \in \mathcal{V}_h\,. \tag{45}$$

Then, (42) is clearly equivalent to the following nonlinear differential system in $\mathbb{R}^{N_h}$ (where $\mathbf{U}_h(t)$ is the vector of the dergrees of freedom of $\mathbf{u}_h(t)$)

$$\mathbf{M}_h \frac{d^2\mathbf{U}_h}{dt^2} + \mathbf{D}_h^*\big(\,\nabla H(\mathbf{D}_h \mathbf{U}_h)\big) = 0 \tag{46}$$

The effective implementation (after time discretization - see the next paragraph) inevitably requires to be able to compute the integrals in the right hand sides of (43), (45). For the nonlinear part (45), it is not possible to compute exactly these integrals - except for very particular H. This is the case for instance with the exact string model. That is why these integrals will be evaluated approximately, which will lead to the following new definitions for $\mathbf{M}_h$ and $\mathbf{D}_h^*\big(\,\nabla H(\mathbf{D}_h)\,\big)$:

$$\left\{\begin{aligned} &\Big(M_h \mathbf{U}_h, \mathbf{V}_h\Big)_h = \oint_\Omega^h \mathbf{u}_h.\mathbf{v}_h, \quad \forall\, \mathbf{v}_h \in \mathcal{V}_h\,,\\ &\Big(\mathbf{D}_h^*\big(\,\nabla H(\mathbf{D}_h \mathbf{U}_h)\,\big), \mathbf{V}_h\,\Big)_h = \oint_\Omega^h \nabla H(\partial_x \mathbf{u}_h)\cdot \partial_x \mathbf{v}_h, \quad \forall\, \mathbf{v}_h \in \mathcal{V}_h\,. \end{aligned}\right. \tag{47}$$

where the linear form $f \mapsto \displaystyle\oint_\Omega^h f$ is an approximate integral

$$\oint_\Omega^h f \simeq \int_\Omega f \quad \text{for small } h. \tag{48}$$

In practice, this approximate integral will be constructed, in the context of finite elements, by decomposing the global integral as the sum of integrals along the segments of the finite element mesh and using inside each segment a given quadrature rule. In practice, as a result, the integral becomes exact as soon as f is piecewise (according to the mesh) polynomial of a certain degree. In particular the calculation of the mass matrix $\mathbf{M}_h$ may be exact in this degree is large enough since, contrary to $\nabla H(\partial_x \mathbf{u}_h)\cdot \partial_x \mathbf{v}_h$, the product $\mathbf{u}_h.\mathbf{v}_h$ is piecewise polynomial.

An important property is required, namely

$$f \geq 0 \quad \Longrightarrow \quad \oint_{\Omega}^{h} f \geq 0, \tag{49}$$

which will be achieved in the finite element context by using quadrature formulas with positive quadrature weights.

Remark 2.1 *In practice, for appropriate quadrature formulas adapted to the finite element space $\mathcal{V}_h$, the positivity of the quadrature weights induces a stronger property, namely the existence of $\gamma > 0$ such that*

$$\forall \mathbf{v}_h \in \mathcal{V}_h, \quad \oint_{\Omega}^{h} |\mathbf{v}_h|^2 \geq \gamma \int_{\Omega} |\mathbf{v}_h|^2.$$

On the other hand, one can choose a quadrature rule that makes the mass matrix become diagonal. This is called mass lumping and can lead to explicit schemes (see section 10.4 pages 305 to 313 of [22] for a mathematical approach).

For a smooth enough function H, typically $H \in C^2(\mathbb{R})$, the existence and uniqueness of a local (in a maximum time interval $[0, T_h[$) solution $\mathbf{u}_h$ of (42) is a direct and easy consequence of standard theorems from the theory of ordinary differential equations [20], with the regularity:

$$\mathbf{u}_h \in C^2(0, T_h; \mathcal{V}_h).$$

Our next result allows us to show that the solution is for each h global in time ($T_h = +\infty$) and provides H^1 stability estimates.

Theorem 2.1 *The scheme (42) preserves a semi discrete energy, ie the solution $\mathbf{u}_h$ of the scheme verifies:*

$$\frac{d}{dt} E_h(t) = 0, \quad \textit{with} \quad E_h(t) = \frac{1}{2} \oint_{\Omega}^{h} \left|\partial_t \mathbf{u}_h\right|^2 + \oint_{\Omega}^{h} H(\partial_x \mathbf{u}_h).$$

Proof. This property comes directly from the variational formulation, taking $\mathbf{v}_h = \partial_t \mathbf{u}_h$ which belongs to $\mathcal{V}_h$. Then we have:

$$\oint_{\Omega}^{h} \left(\partial_t^2 \mathbf{u}_h\right).\left(\partial_t \mathbf{u}_h\right) + \oint_{\Omega}^{h} \nabla H(\partial_x \mathbf{u}_h).\partial_x \left(\partial_t \mathbf{u}_h\right) = 0$$

that can be rewritten

$$\oint_{\Omega}^{h} \partial_t \left(\frac{1}{2} \left|\partial_t \mathbf{u}_h\right|^2\right) + \oint_{\Omega}^{h} \partial_t \left(H(\partial_x \mathbf{u}_h)\right) = 0,$$

which leads to the result.

□

Theorem 2.2 *Let us assume hypothesis (1.2.1), i.e. that there exists $K > 0$ such that:*

$$\forall \mathbf{v} \in \mathbb{R}^N, \ H(\mathbf{v}) \geq K|\mathbf{v}|^2 \tag{50}$$

then, there exists $C > 0$ such that:

$$||\mathbf{u}_h(t)||^2_{H^1_h} \leq C \, E_h(0), \qquad \forall t \geq 0$$

Remark 2.2 *The notation $|| \cdot ||_{H^1_h}$ refers to the approximate integral $\displaystyle\oint_{\Omega}^{h}$.*

Proof.

$$\begin{aligned}
& \oint_\Omega^h \left\{ \frac{1}{2} \left| \frac{\partial \mathbf{u}_h}{\partial t} \right|^2 + H(\frac{\partial \mathbf{u}_h}{\partial x}) \right\} dx = E_h(0), \\
\Rightarrow \quad & \oint_\Omega^h H(\frac{\partial \mathbf{u}_h}{\partial x})\, dx = E_h(0) - \oint_\Omega^h \frac{1}{2} \left| \frac{\partial \mathbf{u}_h}{\partial t} \right|^2 dx, \\
\Rightarrow \quad & K \oint_\Omega^h \left| \frac{\partial \mathbf{u}_h}{\partial x} \right|^2 \leq \oint_\Omega^h H(\frac{\partial \mathbf{u}_h}{\partial x})\, dx \leq E_h(0), \qquad \text{using the hypothesis (49) and (50)} \\
\Rightarrow \quad & \left|\left| \frac{\partial \mathbf{u}_h}{\partial x} \right|\right|^2_{L^2_h} \leq \frac{E_h(0)}{K}.
\end{aligned}$$

Dirichlet boundary condition allows us to use a discrete Poincaré's inequality, and to conclude:

$$||\mathbf{u}_h||^2_{H^1_h} \leq C\, E_h(0).$$

□

In this paragraph we have written the variational formulation of our system of equations (13) which leads to (40), and (42) in its semi discrete form. This semi discrete form requires the calculation of integrals of nonlinear functions, we will use approximate integration (quadrature rules on segments) to tackle this issue. This semi discretization leads naturally to a semi discrete energy preservation (theorem 2.1) and to H^1_h stability of the semi discrete solution (theorem 2.2), providing hypothesis on the potential energy (hypothesis (1.2.1) and (50)), and on the approximate integrals (49).

2.2 Time discretization : construction of energy preserving schemes

As announced previously, we investigate the question of finding finite difference schemes that preserve rigorously a discrete energy. Such schemes are well known in the linear case, which corresponds to

$$H(\mathbf{v}) = \frac{1}{2} \mathbf{A}\mathbf{v} \cdot \mathbf{v}, \quad (\Longrightarrow \quad \nabla H(\mathbf{v}) = \mathbf{A}\mathbf{v}) \tag{51}$$

that is to say to the linear hyperbolic system

$$\partial_t^2 \mathbf{u} - \mathbf{A}\, \partial_{xx}^2 \mathbf{u} = 0 \tag{52}$$

and its corresponding semi-discrete version, using the notation of the previous paragraph

$$\frac{d^2}{dt^2} \left[\oint_\Omega^h \mathbf{u}_h \cdot \mathbf{v}_h \right] + \oint_\Omega^h \mathbf{A}\, \partial_x \mathbf{u}_h \cdot \partial_x \mathbf{v}_h = 0, \quad \forall\, \mathbf{v}_h \in \mathcal{V}_h\,. \tag{53}$$

which preserves the quadratic discrete energy

$$E_h(t) = \frac{1}{2} \oint_\Omega^h |\partial_t \mathbf{u}_h|^2 + \frac{1}{2} \oint_\Omega^h \mathbf{A}\, \partial_x \mathbf{u}_h \cdot \partial_x \mathbf{u}_h \tag{54}$$

In this case, there is a natural class of energy preserving schemes, called the $\theta-$schemes, where $\theta \in [0, 1/2]$ is an averaging parameter. Those schemes belong to the more general class of Newmark schemes that also contain dissipative schemes. Using a constant time step Δt and denoting by $\mathbf{u}_h^n$ the approximation of $\mathbf{u}_h(t^n)$, this scheme is, in its variational form:

$$\oint_\Omega^h \frac{\mathbf{u}_h^{n+1} - 2\mathbf{u}_h^n + \mathbf{u}_h^{n-1}}{\Delta t^2} \cdot \mathbf{v}_h + \oint_\Omega^h \mathbf{A}\, \partial_x \big(\theta \mathbf{u}_h^{n+1} + (1-2\theta)\mathbf{u}_h^n + \theta \mathbf{u}_h^{n-1} \big) \cdot \partial_x \mathbf{v}_h = 0, \quad \forall\, \mathbf{v}_h \in \mathcal{V}_h\,. \tag{55}$$

The solution of this scheme satisfies the conservation of a discrete energy

$$E_h^{n+\frac{1}{2}} = E_h^{n-\frac{1}{2}} \tag{56}$$

where the discrete energy $E_h^{n+\frac{1}{2}}$ corresponding to time $t^{n+\frac{1}{2}} = (n+\frac{1}{2})\Delta t$ is given by

$$\left| \begin{array}{rcl} E_h^{n+\frac{1}{2}} & = & \displaystyle \frac{1}{2} \oint_\Omega^h \Big| \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\Delta t} \Big|^2 + \frac{1}{2} \oint_\Omega^h \mathbf{A}\, \partial_x \big(\mathbf{u}_h^{n+1/2}\big) \cdot \partial_x \big(\mathbf{u}_h^{n+1/2}\big) \\ & + & \displaystyle \frac{1}{2} \big(\theta - \frac{1}{4}\big)\, \Delta t^2 \oint_\Omega^h \mathbf{A}\, \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\Delta t} \cdot \frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\Delta t} \end{array} \right. \tag{57}$$

where we have set $\mathbf{u}_h^{n+1/2} := \dfrac{\mathbf{u}_h^{n+1} + \mathbf{u}_h^n}{2}$.

The identity (56) is easily derived by taking $v_h = \big(\mathbf{u}_h^{n+1} - \mathbf{u}_h^{n-1}\big)/2\Delta t$ as a test function in (55).

The conservation of the energy $E_h^{n+\frac{1}{2}}$ automatically provides the stability of the scheme when $\theta \geq 1/4$ since $E_h^{n+\frac{1}{2}}$ is always positive.

When $\theta < 1/4$, the scheme is stable under the stability condition

$$(1-4\theta)\, \frac{\Delta t^2}{4} \sup_{\mathbf{v}_h \in \mathcal{V}_h} \left[\frac{\displaystyle \oint_\Omega^h \mathbf{A}\, \partial_x \mathbf{v}_h \cdot \partial_x \mathbf{v}_h}{\displaystyle \oint_\Omega^h |\mathbf{v}_h|^2} \right] \leq 1 \tag{58}$$

which is nothing but the condition that ensures the positivity of $E_h^{n+\frac{1}{2}}$.

When $\theta = 0$, one gets the well-known leap-frog scheme - or explicit scheme - which is explicit in practice when one achieves mass lumping (see remark 2.1).

Our objective is in some sense to generalize the θ-scheme to the nonlinear case, with as main objective the preservation of a discrete energy guaranteeing the stability of the scheme.

Anticipating some of our results, we shall see that our wish to preserve a discrete energy will lead to consider implicit schemes. In the context that has motivated the present work, i.e. the numerical modeling of a piano, this is by no means a real constraint : the full piano model couples the vibrations of the string (a 1D model) with the vibrations of the soundboard (a 2D phenomenon) and with the sound radiation (a 2D phenomenon). Thus, in the full computational code, the time devoted to the string itself is a very small percentage of the total computational cost and we are ready to pay the possible increase of computational time (due to implicitness) for the string if it gives us more flexibility and robustness for the coupled model. The implicit nature of the scheme will lead to unconditional stability of the scheme. As a consequence, we shall have no mathematical constraint on the choice of the time step, which is a pleasant property for the coupled problem. Moreover, the fact that we have a stability proof by energy will also allow us to treat the stability of the coupled problem, which is not a priori an easy task.

The problematic of energy preserving schemes in the nonlinear case is obviously not new. In the simpler context of systems of nonlinear ordinary differential equations, several contributions are due to Mickens [28] and Kevrekidis [24] in the case of polynomial non linearities, to Greenspan [17] in the case of general scalar nonlinearity (see also Chin and Qint [7] for the three body problem and Gonzalez and Simo [16] for a particle in a potential). Several authors (see for instance

[37, 11, 12, 5]) tackle the case of the scalar ($N = 1$) nonlinear wave equation. We shall come back to this case in section 2.2.1. The case of systems ($N > 1$) appears to be more delicate and has apparently retained much less attention, at least in the general case. We can mention the work of Gonzalez [15] for nonlinear elasticity and Kane, Marsden and Ortiz [23] in a general lagrangian case. However, for studying the vibrations of a non linear string, S. Bilbao in [4] has proposed for the Bank-Sujbert model [2] a numerical scheme which preserves a discrete energy under a suitable stability condition (that coïncides with the stability condition for the explicit scheme for the linear model (58)). The way this scheme has been derived is rather mysterious but clearly exploits the polynomial nature of the function H (as in [28, 24]), which makes its extension to the general case difficult. This scheme is implicit but the the polynomial nature of H gives a special structure to the problem that can be exploited from the computational point of view.

Of course the problematic of energy preserving schemes is close to the problematic of symplectic schemes [34] for the discretization of hamiltonian differential equations, whose purpose is to preserve other invariants as in the continuous problems. These invariants are of more geometrical nature and linked to the preservation of symmetries of the system. In general, such schemes cannot preserve a discrete energy (see [38]) but can succeed in "almost preserving" such an energy over large times [18, 19, 32]. Several authors have lead a comparison between symplectic and preserving schemes, and found the latter "more accurate" (see [16, 15, 10])

In what follows, we are going to investigate a class of three point schemes for the time discretization of (42). These schemes have the same type of structure that the θ-schemes and include all the linear schemes. These schemes will be based on a function that we shall call the "approximate gradient":

$$\left|\begin{array}{llll} \mathbb{V}H: & \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N & \longrightarrow & \mathbb{R}^N \\ & & & \\ & (\mathbf{u},\mathbf{v},\mathbf{w}) & \longrightarrow & \mathbb{V}H(\mathbf{u},\mathbf{v},\mathbf{w}) \end{array}\right. \tag{59}$$

that should satisfy the consistency condition

$$\forall \mathbf{v} \in \mathbb{R}^N, \quad \mathbb{V}H(\mathbf{v},\mathbf{v},\mathbf{v}) = \nabla H(\mathbf{v}). \tag{60}$$

Using such an approximate gradient, the fully discrete version of (53) is

$$\oint_\Omega^h \frac{\mathbf{u}_h^{n+1} - 2\mathbf{u}_h^n + \mathbf{u}_h^{n-1}}{\Delta t^2} \cdot \mathbf{v}_h + \oint_\Omega^h \mathbb{V}H(\partial_x \mathbf{u}_h^{n+1}, \partial_x \mathbf{u}_h^n, \partial_x \mathbf{u}_h^{n-1}) \cdot \partial_x \mathbf{v}_h = 0, \quad \forall\, \mathbf{v}_h \in \mathcal{V}_h\,. \tag{61}$$

For the sequel, we shall assume that $\mathbb{V}H$ is a "smooth enough" function. Note that:

- One obtains an explicit scheme (provided mass lumping) as soon as

$$\mathbb{V}H(\mathbf{u},\mathbf{v},\mathbf{w}) \text{ is independent of } \mathbf{u}. \tag{62}$$

- One obtains a scheme which is reversible in time and second order accurate (see remark 2.3) if and only if

$$\mathbb{V}H(\mathbf{u},\mathbf{v},\mathbf{w}) = \mathbb{V}H(\mathbf{w},\mathbf{v},\mathbf{u}), \quad \forall\, (\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N. \tag{63}$$

Remark 2.3 *To check the second order accuracy of the time approximation, we introduce the truncation error defined as the linear form*

$$\mathcal{E}(\mathbf{v}_h) = \oint_\Omega^h \frac{\mathbf{U}_h^{n+1} - 2\mathbf{U}_h^n + \mathbf{U}_h^{n-1}}{\Delta t^2} \cdot \mathbf{v}_h + \oint_\Omega^h \mathbb{V}H(\partial_x \mathbf{u}_h^{n+1}, \partial_x \mathbf{U}_h^n, \partial_x \mathbf{U}_h^{n-1}) \cdot \partial_x \mathbf{v}_h \tag{64}$$

with $U_h^n := \mathbf{u}_h(t)$ and $\mathbf{u}_h(\cdot)$ is the solution of (42), which is a smooth function of time. By a Taylor expansion we have

$$\frac{\mathbf{U}_h^{n+1} - 2\mathbf{U}_h^n + \mathbf{U}_h^{n-1}}{\Delta t^2} = \frac{d^2\mathbf{u}_h}{dt^2}(t^n) + O(\Delta t^2) \tag{65}$$

On the other hand, denoting $D_1\nabla H(\mathbf{u},\mathbf{v},\mathbf{w})$ *(resp.* $D_3\nabla H(\mathbf{u},\mathbf{v},\mathbf{w})$*) the differential of the application* $\mathbf{u}\mapsto\nabla H(\mathbf{u},\mathbf{v},\mathbf{w})$ *(resp.*$\mathbf{w}\mapsto\nabla H(\mathbf{u},\mathbf{v},\mathbf{w})$*), we have*

$$\begin{aligned}\nabla H(\partial_x\mathbf{U}_h^{n+1},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n-1}) &= \nabla H(\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n})\\ &+ D_1\nabla H(\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n})\,\partial_x\big(\mathbf{U}_h^{n+1}-\mathbf{U}_h^{n}\big)\\ &+ D_3\nabla H(\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n})\,\partial_x\big(\mathbf{U}_h^{n-1}-\mathbf{U}_h^{n}\big)+O(\Delta t^2)\end{aligned}$$

that is to say, using the consistency condition (60)

$$\begin{aligned}\nabla H(\partial_x\mathbf{U}_h^{n+1},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n-1}) &= \nabla H(\mathbf{U}_h^{n})\\ &+ D_1\nabla H(\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n})\,\partial_x\big(\mathbf{U}_h^{n+1}-\mathbf{U}_h^{n}\big)\\ &+ D_3\nabla H(\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n})\,\partial_x\big(\mathbf{U}_h^{n-1}-\mathbf{U}_h^{n}\big)+O(\Delta t^2)\end{aligned}$$

Differentiating with respect to $\mathbf{u}$ *the symmetry condition (63), we get*

$$D_1\nabla H(\mathbf{u},\mathbf{v},\mathbf{w}) = D_3\nabla H(\mathbf{w},\mathbf{v},\mathbf{u}),\quad \forall\,(\mathbf{u},\mathbf{v},\mathbf{w})\in\mathbb{R}^N\times\mathbb{R}^N\times\mathbb{R}^N.$$

so that we can write

$$\begin{aligned}\nabla H(\partial_x\mathbf{U}_h^{n+1},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n-1}) &= \nabla H(\partial_x\mathbf{U}_h^{n})+O(\Delta t^2)\\ &+ D_1\nabla H(\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n})\,\big(\partial_x\mathbf{U}_h^{n+1}-2\partial_x\mathbf{U}_h^{n}+\partial_x\mathbf{U}_h^{n-1}\big)\end{aligned}$$

that is to say , since $\big(\partial_x\mathbf{U}_h^{n+1}-2\partial_x\mathbf{U}_h^{n}+\partial_x\mathbf{U}_h^{n-1}\big)=O(\Delta t^2)$

$$\nabla H(\partial_x\mathbf{U}_h^{n+1},\partial_x\mathbf{U}_h^{n},\partial_x\mathbf{U}_h^{n-1}) = \nabla H(\partial_x\mathbf{U}_h^{n})+O(\Delta t^2) \tag{66}$$

It suffices to substitute (65) and (66) into (64) and to use the fact that $\mathbf{u}_h(\cdot)$ *is solution of (42) to conclude that*

$$\mathcal{E}(\mathbf{v}_h)=O(\Delta t^2).$$

To investigate the preservation of a discrete energy, we choose $v_h=\big(\mathbf{u}_h^{n+1}-\mathbf{u}_h^{n-1}\big)/2\Delta t$ as a test function in (61) (as in the linear case) and obtain

$$\left|\begin{aligned}&\frac{1}{\Delta t}\Big\{\frac12\oint_\Omega^h\Big|\frac{\mathbf{u}_h^{n+1}-\mathbf{u}_h^{n}}{\Delta t}\Big|^2-\frac12\oint_\Omega^h\Big|\frac{\mathbf{u}_h^{n}-\mathbf{u}_h^{n-1}}{\Delta t}\Big|^2\Big\}\\ &+\oint_\Omega^h\nabla H(\partial_x\mathbf{u}_h^{n+1},\partial_x\mathbf{u}_h^{n},\partial_x\mathbf{u}_h^{n-1})\cdot\partial_x\Big(\frac{\mathbf{u}_h^{n+1}-\mathbf{u}_h^{n-1}}{2\Delta t}\Big)=0.\end{aligned}\right. \tag{67}$$

This identity suggests us the following

Definition 2.1 *The function* ∇H *is called "conservative" (we shall also say that the corresponding scheme is "conservative" or "energy preserving"- cf lemma 2.2.1) if and only if there exists a scalar function (a discrete potential energy)*

$$\left|\begin{array}{llll}\mathbb{H}: & \mathbb{R}^N\times\mathbb{R}^N & \longrightarrow & \mathbb{R}\\ & (\mathbf{u},\mathbf{v}) & \longrightarrow & \mathbb{H}(\mathbf{u},\mathbf{v})\end{array}\right. \tag{68}$$

such that

$$\forall\,(\mathbf{u},\mathbf{v},\mathbf{w})\in\mathbb{R}^N\times\mathbb{R}^N\times\mathbb{R}^N,\quad \nabla H(\mathbf{u},\mathbf{v},\mathbf{w})\cdot\frac{(\mathbf{u}-\mathbf{w})}{2}=\mathbb{H}(\mathbf{u},\mathbf{v})-\mathbb{H}(\mathbf{v},\mathbf{w}). \tag{69}$$

Remark 2.4 *Note that (69) implies in particular the symmetry of* $\mathbb{H}$*:*

$$\forall\,(\mathbf{u},\mathbf{v}) \in \mathbb{R}^N \times \mathbb{R}^N, \quad \mathbb{H}(\mathbf{u},\mathbf{v}) = \mathbb{H}(\mathbf{v},\mathbf{u}). \tag{70}$$

Of course, for consistency reasons, the discrete potential energy $\mathbb{H}$ should also satisfy (in agreement with (60)

$$\forall \mathbf{v} \in \mathbb{R}^N, \quad \mathbb{H}(\mathbf{v},\mathbf{v}) = H(\mathbf{v}). \tag{71}$$

Assuming that ∇H is conservative, we deduce from (69) that

$$\left| \begin{array}{l} \nabla H(\partial_x \mathbf{u}_h^{n+1}, \partial_x \mathbf{u}_h^{n}, \partial_x \mathbf{u}_h^{n-1}) \,\cdot\, \partial_x \big(\dfrac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^{n-1}}{2\Delta t}\big) = \\[2ex] \qquad = \dfrac{1}{\Delta t} \Big\{\, \mathbb{H}(\partial_x \mathbf{u}_h^{n+1}, \partial_x \mathbf{u}_h^{n}) - \mathbb{H}(\mathbf{u}_h^{n}, \partial_x \mathbf{u}_h^{n-1}) \,\Big\} \end{array} \right. \tag{72}$$

which is a discrete equivalent of the derivation rule for composed functions:

$$\frac{\partial}{\partial t} H(\partial_x \mathbf{u}) = \nabla H(\partial_x \mathbf{u}) \cdot \partial^2_{xt} \mathbf{u}. \tag{73}$$

Joining (72) to (67) leads to the:

Lemma 2.2.1 *If u_h^n is a solution of (61), with ∇H conservative in the sense of definition 2.1, it satisfies the energy conservation property:*

$$E_h^{n+\frac{1}{2}} = E_h^{n-\frac{1}{2}} \tag{74}$$

where the discrete energy $E_h^{n+\frac{1}{2}}$ corresponding to time $t^{n+\frac{1}{2}} = (n+\frac{1}{2})\Delta t$ is given by

$$E_h^{n+\frac{1}{2}} = \frac{1}{2} \oint_\Omega^h \big|\frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^{n}}{\Delta t}\big|^2 + \oint_\Omega^h \mathbb{H}\big(\,\partial_x \mathbf{u}_h^{n+1}, \partial_x \mathbf{u}_h^{n}\big) \tag{75}$$

where $\mathbb{H}$ is the discrete potential energy associated to ∇H (see (69)).

An immediate consequence (we omit the details of the proof) of this lemma is the

Corollary 2.2.1 *If the discrete potential energy $\mathbb{H}$ is positive*

$$\forall (\mathbf{u},\mathbf{v}) \in \mathbb{R}^N, \quad \mathbb{H}(\mathbf{u},\mathbf{v}) \geq 0, \tag{76}$$

then the scheme is unconditionally L^2-stable in the sense that the L^2 norm in space of any solution u_h^n is uniformly (with respect to h and Δt) bounded.

Remark 2.5 *In the linear case, i.e. for $H(\mathbf{v}) = \frac{1}{2}; \mathbf{A}\,\mathbf{v}\cdot\mathbf{v}$, the θ-scheme corresponds to*

$$\left| \begin{array}{l} \nabla H(\mathbf{u},\mathbf{v},\mathbf{w}) = \mathbf{A}\,\big(\theta \mathbf{u} + (1-2\theta)\mathbf{v} + \theta \mathbf{w}\big) \\[2ex] \mathbb{H}(\mathbf{u},\mathbf{v}) = \dfrac{1}{2}\,\mathbf{A}\,\dfrac{\mathbf{u}+\mathbf{v}}{2} \cdot \dfrac{\mathbf{u}+\mathbf{v}}{2} + \dfrac{1}{2}\,(\theta - \dfrac{1}{4})\mathbf{A}\,(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}). \end{array} \right.$$

and the positivity property (76) is unconditionally achieved if and only if $\theta \geq \frac{1}{4}$.

Before investigating more elaborated discretizations, let us consider the case of the most naïve scheme to discretize (42), which is the most natural extension to the non linear case of the explicit leap frog scheme ($\theta = 0$) for the linear case:

$$\oint_\Omega^h \frac{\mathbf{u}_h^{n+1} - 2\mathbf{u}_h^{n} + \mathbf{u}_h^{n-1}}{\Delta t^2} \cdot \mathbf{v}_h + \oint_\Omega^h \nabla H(\partial_x \mathbf{u}_h^{n}) \cdot \partial_x \mathbf{v}_h = 0, \quad \forall\ \mathbf{v}_h \in \mathcal{V}_h\,. \tag{77}$$

which corresponds to

$$\nabla H(\mathbf{u},\mathbf{v},\mathbf{w}) = \nabla H(\mathbf{v}). \tag{78}$$

Lemma 2.2.2 *The explicit scheme (77) is conservative in the sense of the definition 2.1 if and only if the original equation is linear.*

Proof. According to definition 2.1, we look for a discrete potential energy $\mathbb{H}(\mathbf{u}.\mathbf{v})$ such that

$$\forall\,(\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N, \quad \mathbb{H}(\mathbf{u},\mathbf{v}) - \mathbb{H}(\mathbf{v},\mathbf{w}) = F(\mathbf{v}) \cdot \frac{(\mathbf{u}-\mathbf{w})}{2}, \text{ with } F(\mathbf{v}) := \nabla H(\mathbf{v}). \quad (79)$$

Our objective is to show that if $\mathbb{H}(\mathbf{u},\mathbf{v})$ exists, $F(\mathbf{v}) = \nabla H(\mathbf{v})$ is necessarily linear in $\mathbf{v}$. In what follows we shall denote $\nabla_1 \mathbb{H}(\mathbf{u},\mathbf{v})$ (resp $\nabla_2 \mathbb{H}(\mathbf{u},\mathbf{v})$) the gradient of the function $\mathbf{u} \mapsto \mathbb{H}(\mathbf{u},\mathbf{v})$ (respectively $\mathbf{v} \mapsto \mathbb{H}(\mathbf{u},\mathbf{v})$).

First, differentiating (79) with respect to $\mathbf{u}$ leads to the identity

$$\nabla_1 \mathbb{H}(\mathbf{v},\mathbf{w}) = \frac{1}{2} F(\mathbf{v}), \quad \forall\,(\mathbf{v},\mathbf{w}) \in \mathbb{R}^N \times \mathbb{R}^N. \quad (80)$$

Next, differentiating (79) with respect to $\mathbf{v}$ leads to ($DF(\mathbf{u}) \in \mathcal{L}(\mathbb{R}^N)$ denotes the differential of $F(\mathbf{u})$ and $DF(\mathbf{u})^*$ its adjoint with respect the usual inner product in $\mathbb{R}^N$)

$$\nabla_2 \mathbb{H}(\mathbf{u},\mathbf{v}) - \nabla_1 \mathbb{H}(\mathbf{v},\mathbf{w}) = \frac{1}{2} DF(\mathbf{v})^*(\mathbf{u}-\mathbf{w}), \quad \forall\,(\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N, \quad (81)$$

which can be rewritten, thanks to (80)

$$\nabla_2 \mathbb{H}(\mathbf{u},\mathbf{v}) = \frac{1}{2} F(\mathbf{w}) + \frac{1}{2} DF(\mathbf{v})^*(\mathbf{u}-\mathbf{w}),, \quad \forall\,(\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N. \quad (82)$$

Finally, we differentiate (82) with respect to $\mathbf{w}$ to obtain

$$0 = \frac{1}{2} DF(\mathbf{w})^* - \frac{1}{2} DF(\mathbf{v})^*), \quad \forall\,(\mathbf{v},\mathbf{w}) \in \mathbb{R}^N \times \mathbb{R}^N. \quad (83)$$

This means that $DF(\mathbf{v})$ is constant in $\mathbf{v}$, i.e. that F is linear, which achieves the proof.

□

We have in fact a more general result.

Lemma 2.2.3 *Let us consider a scheme of the form (61) that is explicit, i. e. ∇H is independent of $\mathbf{u}$, and consistent. It is conservative in the sense of the definition 2.1 if and only if ∇H is linear and $\nabla H(\mathbf{u},\mathbf{v},\mathbf{w}) = \nabla H(\mathbf{v})$.*

Proof. Assume that there exists a discrete potential energy $\mathbb{H}(\mathbf{u}.\mathbf{v})$ such that

$$\forall\,(\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N, \quad \mathbb{H}(\mathbf{u},\mathbf{v}) - \mathbb{H}(\mathbf{v},\mathbf{w}) = F(\mathbf{v},\mathbf{w}) \cdot \frac{(\mathbf{u}-\mathbf{w})}{2},. \quad (84)$$

with $F(\mathbf{v}) := \nabla H(\mathbf{v},\mathbf{w})$.

Differentiating twice (84) once with respect to u the other with respect to $\mathbf{w}$, we get

$$D_2 F(\mathbf{v},\mathbf{w}) = 0$$

where $D_2 F(\mathbf{v},\mathbf{w})$ is the differential of the function $\mathbf{w} \mapsto F(\mathbf{v},\mathbf{w})$. This means that $F(\mathbf{v},\mathbf{w})$ does not depend on $\mathbf{w}$, i. e.

$$F(\mathbf{v},\mathbf{w}) \equiv F(\mathbf{v}).$$

The consistency condition then implies $F(\mathbf{v}) = \nabla H(\mathbf{v})$ and we can then use lemma 2.2.2 to conclude.

□

This last lemma shows that, except in the linear case, conservativity implies implicitness.

2.2.1 The case of the scalar nonlinear wave equation (N=1)

We consider the scalar wave equation (N=1), in which case we can write without any ambiguity (we omit the index h for simplicity)

$$\mathbf{u} \equiv u_1, \tag{85}$$

and (35) simply becomes

$$\partial_t^2 u_1 - \partial_x \big[F(\partial_x u_1) \big] = 0, \quad F := H' \quad (\equiv \nabla H) \tag{86}$$

and the scheme (61) writes

$$\oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_h^{n-1}}{\Delta t^2} \cdot v_1 + \oint_\Omega^h \nabla H(\partial_x u_1^{n+1}, \partial_x u_1^n, \partial_x u_1^{n-1}) \cdot \partial_x v_1 = 0, \quad \forall\, v_1 \in \mathcal{V}_h\,. \tag{87}$$

The conservativity condition (69) is simply

$$\forall\,(u_1, v_1, w_1) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}, \quad \nabla H(u_1, v_1, w_1) \cdot \frac{(u_1 - w_1)}{2} = \mathbb{H}(u_1, v_1) - \mathbb{H}(v_1, w_1). \tag{88}$$

We notice that, given a discrete energy function $\mathbb{H}(u_1, v_1) : \mathbb{R}^2 \mapsto \mathbb{R}$, satisfying the symmetry condition (that me know to be necessary - cf remark 2.4):

$$\forall\,(u_1, v_1) \in \mathbb{R} \times \mathbb{R}, \quad \mathbb{H}(u_1, v_1) = \mathbb{H}(v_1, u_1). \tag{89}$$

(88) determines completely $\nabla H(u_1, v_1, w_1)$ (this is due to the fact that is $\nabla H(u_1, v_1, w_1)$ real valued which holds only when $N = 1$) as (note that we use the symmetry of $\mathbb{H}$ and that ∇H is smooth):

$$\nabla H(u_1, v_1, w_1) = \begin{cases} \dfrac{\mathbb{H}(u_1, v_1) - \mathbb{H}(w_1, v_1)}{u_1 - w_1} & \text{if } u_1 \neq w_1, \\[2ex] \dfrac{\partial \mathbb{H}}{\partial u_1}(u_1, v_1) \equiv \dfrac{\partial \mathbb{H}}{\partial v_1}(u_1, v_1) & \text{if } u_1 = w_1. \end{cases} \tag{90}$$

In fact given any positive symmetric function $\mathbb{H}(u_1, v_1)$ satisfying the consistency condition

$$\forall v_1 \in \mathbb{R}, \quad \mathbb{H}(v_1, v_1) = H(v_1), \tag{91}$$

the choice of $\nabla H(u_1, v_1, w_1)$ given by (90) provides a consistent, energy preserving numerical scheme.

The rest is simply a question of choice of a positive function $\mathbb{H}$ satisfying both symmetry (89) and consistency (91). The two simplest choices (in our opinion) are

$$\mathbb{H}(u_1, v_1) = \frac{1}{2} \Big\{ H(u_1) + H(v_1) \Big\} \tag{92}$$

$$\mathbb{H}(u_1, v_1) = H\Big(\frac{u_1 + v_1}{2} \Big) \tag{93}$$

The choice (92) leads to the scheme

$$\oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_h^{n-1}}{\Delta t^2} \cdot v_1 + \oint_\Omega^h \frac{H(\partial_x u_1^{n+1}) - H(\partial_x u_1^{n-1})}{\partial_x u_1^{n+1} - \partial_x u_1^{n-1}} \cdot \partial_x v_1 = 0, \quad \forall\, v_1 \in \mathcal{V}_h\,. \tag{94}$$

while the choice (93) leads to the scheme

$$\oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_h^{n-1}}{\Delta t^2} \cdot v_1 + \oint_\Omega^h \frac{H(\partial_x u_1^{n+1/2}) - H(\partial_x u_1^{n-1/2})}{\partial_x u_1^{n+1/2} - \partial_x u_1^{n-1/2}} \cdot \partial_x v_1 = 0, \quad \forall\, v_1 \in \mathcal{V}_h\,. \tag{95}$$

where by convention $u_1^{n+1/2} = \dfrac{u_1^{n+1} + u_1^n}{2}$.

The reader will easily check that in the linear case, the scheme (94) gives the θ-scheme with $\theta = 1/2$ while the scheme (95) gives the θ-scheme with $\theta = 1/4$. Other θ-schemes can be recovered in the linear case by choosing for the discrete energy $\mathbb{H}$ an appropriate linear combination of the two functions (94) and (95).

Remark 2.6 *The schemes are not rigorously defined because of the presence of the denominators. To be more rigorous, we should introduce, for any function of one variable $\Phi : \mathbb{R} \mapsto \mathbb{R}$, the function of 2 variables:*

$$\delta\Phi(u_1, w_1) = \begin{cases} \dfrac{\Phi(u_1) - \Phi(w_1)}{u_1 - w_1} & \textit{if } u_1 \neq w_1, \\ \Phi'(w_1) & \textit{if } u_1 = w_1. \end{cases} \tag{96}$$

and rewrite (94) and (95) as respectively

$$\oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_h^{n-1}}{\Delta t^2} \cdot v_1 + \oint_\Omega^h \delta H(\partial_x u_1^{n+1}, \partial_x u_1^{n-1}) \cdot \partial_x v_1 = 0, \quad \forall\, v_1 \in \mathcal{V}_h\,. \tag{97}$$

$$\oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_h^{n-1}}{\Delta t^2} \cdot v_1 + \oint_\Omega^h \delta H(\partial_x u_1^{n+1/2}, u_1^{n-1/2}) \cdot \partial_x v_1 = 0, \quad \forall\, v_1 \in \mathcal{V}_h\,. \tag{98}$$

Remark 2.7 *This form is found in several publications in the scalar case ([12, 24, 17, 37, 35]).*

2.2.2 A class of partially decoupled implicit schemes

We come back to the general case of systems, i.e. $N \geq 1$ and set $\mathbf{u}_h = (u_1, u_2, \cdots, u_N)$. If we look at the equation number ℓ of the system (35), it writes:

$$\partial_t^2 u_\ell - \partial_x \left[\partial_\ell H(\partial_x u_\ell, \partial_x u_{j\neq\ell})\right] = 0 \tag{99}$$

where $\partial_\ell H$ denotes the partial derivative of H with respect to its ℓ^{th} variable and by convention, for any $\mathbf{v} = (v_j) \in \mathbb{R}^N$

$$v_{j\neq c} := (v_1, \cdots, v_{\ell-1}, v_{\ell-1}, \cdots, v_N) \quad \text{ and } \mathbf{v} = (v_j) \equiv (v_\ell, v_{j\neq\ell}). \tag{100}$$

Assuming that $u_{j\neq\ell}$ is known, this is for u_ℓ a 1D equation very similar to (86). Thus, the most natural generalization of the scheme (97) is (with $\mathcal{V}_h = \prod \mathcal{V}_\ell$)

$$\left|\; \begin{aligned} &\oint_\Omega^h \frac{u_\ell^{n+1} - 2u_\ell^n + u_\ell^{n-1}}{\Delta t^2} \cdot v_1 + \oint_\Omega^h \delta_\ell H(\partial_x u_\ell^{n+1}, \partial_x u_\ell^{n-1}; \partial_x u_{j\neq\ell}^n) \cdot \partial_x v_\ell = 0, \\ &\hspace{18em} \forall\, v_\ell \in \mathcal{V}_\ell, \quad \ell = 1, \cdots, N. \end{aligned} \right. \tag{101}$$

where we have introduced a new notation for the multidimensional generalization of (96) : to any scalar function of N variables $\Phi(v_1, \cdots, v_N)$, we associate the function of $N+1$ variables (with notations similar to the previous ones - we omit the details):

$$\delta_\ell\Phi(u_\ell, w_\ell; v_{j\neq\ell}) = \begin{cases} \dfrac{\Phi(u_\ell, v_{j\neq\ell}) - \Phi(w_\ell, v_{j\neq\ell})}{u_\ell - w_\ell} & \text{if } u_\ell \neq w_\ell, \\ \partial_\ell\Phi(w_\ell, v_{j\neq\ell}) & \text{if } u_\ell = w_\ell. \end{cases} \tag{102}$$

which satisfies in particular

$$\delta_\ell\Phi(u_\ell, w_\ell; v_{j\neq\ell})\,(u_\ell - w_\ell) = \Phi(u_\ell, v_{j\neq\ell}) - \Phi(w_\ell, v_{j\neq\ell}) \tag{103}$$

The scheme (101) is clearly of the form (61) with

$$\nabla H(\mathbf{u},\mathbf{v},\mathbf{w}) := \big(\ \delta_\ell H(u_\ell, w_\ell; v_{j\neq\ell})\ \big)_{\ell\ \in[1,\dots,N]}. \tag{104}$$

The question is : is this scheme energy preserving ? When $N = 2$, the answer is clearly yes. It suffices to introduce the discrete energy

$$\mathbb{H}(\mathbf{u},\mathbf{v}) = \frac{1}{2}\Big\{H(u_1,v_2) + H(v_1,u_2)\Big\} \tag{105}$$

which is clearly consistent (cf 71) and to compute that, using (103)

$$\left|\begin{array}{rcl} \nabla H(\mathbf{u},\mathbf{v},\mathbf{w})\cdot\dfrac{(\mathbf{u}-\mathbf{w})}{2} & = & \dfrac{1}{2}\Big\{\ \delta_1 H(u_1,w_1;v_2)\,(u_1-w_1) + \delta_2 H(u_2,w_2;v_1)\,(u_2-w_2)\ \Big\} \\ & = & \dfrac{1}{2}\Big\{\ H(u_1,v_2) - H(w_1,v_2) + H(v_1,u_2) - H(v_1,w_2)\ \Big\} \\ & = & \dfrac{1}{2}\Big\{\ H(u_1,v_2) + H(v_1,u_2)\ \Big\} - \dfrac{1}{2}\Big\{\ H(v_1,w_2) + H(w_1,v_2)\ \Big\} \\ & = & \mathbb{H}(\mathbf{u},\mathbf{v}) - \mathbb{H}(\mathbf{v},\mathbf{w}) \end{array}\right.$$

Unfortunately, this nice property seems limited to the dimension 2, except for very special potential energies H which can be written as the sum of functions of two variables. Let us state a precise result:

Lemma 2.2.4 *Assume $N \geq 2$. The approximate gradient defined by (104) is conservative in the sense of definition 2.1 if and only if*

$$\exists\ \Big\{H_{ij} : \mathbb{R}^2 \mapsto \mathbb{R}, 1 \leq i < j \leq N\Big\} \quad \textit{such that} \quad H(\mathbf{v}) = \sum_{i<j} H_{ij}(v_i,v_j) \tag{106}$$

or equivalently

$$\forall\ i<j<k, \quad \partial^3_{ijk}\ H(\mathbf{v}) = 0. \tag{107}$$

In this case the discrete potential energy $\mathbb{H}$ is given by

$$\mathbb{H}(\mathbf{u},\mathbf{v}) = \frac{1}{2}\sum_{i<j}\ \{H_{ij}(u_i,v_j) + H_{ij}(v_i,u_j)\} \tag{108}$$

Proof. See Appendix A.

□

2.2.3 Construction of fully implicit preserving schemes

Following the same idea as in the previous section, we can first look for schemes of the form:

$$\left|\begin{array}{l} \displaystyle\oint_\Omega^h \frac{u_\ell^{n+1} - 2u_\ell^n + u_\ell^{n-1}}{\Delta t^2}\cdot v_\ell + \oint_\Omega^h \delta_\ell H(\partial_x u_\ell^{n+1}, \partial_x u_\ell^{n-1}; \partial_x [\![u_j^n]\!]^\ell_{j\neq\ell})\cdot \partial_x v_\ell = 0, \\ \hfill \forall\ v_\ell \in \mathcal{V}_\ell, \quad \ell = 1,\cdots,N. \end{array}\right. \tag{109}$$

where $[\![u_j^n]\!]^\ell$ represents some approximation of $u_j(t^n)$, to be determined, using u_j^{n+1}, u_j^n and u_j^{n+1} and that we authorize to depend on ℓ. The decoupled schemes of section 2.2.2 corresponded to $[\![u_j^n]\!] = u_j^n$. We need another choice to ensure conservativity for any n. Let us consider

$$[\![u_j^n]\!]^\ell = u_j^{n+sg(\ell-j)}, \text{ where } sg(\cdot) \text{ is the usual sign function} \tag{110}$$

The interpretation of this choice is the following (the principle is similar to the one of the Gauss-Seidel algorithm for linear systems) : in the l^{th} equation of (115), associated to the component u_ℓ, the other components u_j are evaluated at time t^{n-1} if $j < \ell$ and at time t^{n+1} if $j > \ell$. As an illustration, for $N = 2$ our scheme writes

$$\left\{\begin{array}{l} \displaystyle \oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_1^{n-1}}{\Delta t^2} \cdot v_1 + \oint_\Omega^h \frac{H(\partial_x u_1^{n+1}, \partial_x u_2^{n-1}) - H(\partial_x u_1^{n-1}, \partial_x u_2^{n-1})}{\partial_x u_1^{n+1} - \partial_x u_1^{n-1}} \cdot \partial_x v_1 = 0, \\ \hfill \forall\, v_1 \in \mathcal{V}_1, \\ \displaystyle \oint_\Omega^h \frac{u_2^{n+1} - 2u_2^n + u_2^{n-1}}{\Delta t^2} \cdot v_2 + \oint_\Omega^h \frac{H(\partial_x u_1^{n+1}, \partial_x u_2^{n+1}) - H(\partial_x u_1^{n+1}, \partial_x u_2^{n-1})}{\partial_x u_2^{n+1} - \partial_x u_2^{n-1}} \cdot \partial_x v_2 = 0, \\ \hfill \forall\, v_2 \in \mathcal{V}_2. \end{array}\right. \tag{111}$$

With the choice (110), the scheme (115) enters the class (61) with

$$\Big[\nabla H(\mathbf{u}, \mathbf{v}, \mathbf{w})\Big]_\ell := \big(\ \delta_\ell H(u_\ell, w_\ell; \big[\beta_{j\ell} u_j + (1 - \beta_{j\ell}) w_j\big]_{j \neq \ell})\ \big), \tag{112}$$

where by definition

$$\beta_{j\ell} = 1 \text{ if } j < \ell, \quad 0 \text{ if } j > \ell. \tag{113}$$

Lemma 2.2.5 *The approximate gradient (112) is conservative in the sense of definition 2.1 and the associate discrete potential energy is given by:*

$$\mathbb{H}(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\, \Big\{ H(\mathbf{u}) + H(\mathbf{v}) \Big\} \tag{114}$$

Proof.

$$\nabla H(\mathbf{u}, \mathbf{v}, \mathbf{w}) \cdot \frac{(\mathbf{u} - \mathbf{w})}{2} = \frac{1}{2} \sum_\ell \big(\ \delta_\ell H(u_\ell, w_\ell; \big[\beta_{j\ell} u_j + (1 - \beta_{j\ell}) w_j\big]_{j \neq \ell})\ \big)\, (u_\ell - w_\ell)$$

Using (103) and the definition of $\beta_{j\ell}$, we can write (with some obvious abuse of notation)

$$\nabla H(\mathbf{u}, \mathbf{v}, \mathbf{w}) \cdot \frac{(\mathbf{u} - \mathbf{w})}{2} = \frac{1}{2} \Big[\sum_\ell H(u_1, \cdots, u_\ell, w_{\ell+1}, \cdots w_N) - H(u_1, \cdots, u_{\ell-1}, w_\ell, \cdots w_N) \Big]$$

As this is a telescopic sum, this results into:

$$\nabla H(\mathbf{u}, \mathbf{v}, \mathbf{w}) \cdot \frac{(\mathbf{u} - \mathbf{w})}{2} = \frac{1}{2}\, H(u_1, \cdots, u_N) - \frac{1}{2}\, H(w_1, \cdots, w_N) \equiv \mathbb{H}(\mathbf{u}, \mathbf{v}) - \mathbb{H}(\mathbf{v}, \mathbf{w}).$$

which achieves the proof.

□

Remark 2.8 *The author will notice that, at a given time t^{n+1}, if the equations are solved one by one when increasing the value of ℓ, the calculation of each component can be done in a decoupled way as for the scheme (101): one has to solve a triangular non linear system instead of a diagonal non linear system with (101).*

As emphasized by remark 2.8, the numbering of the equations of (or equally the ranking of the components of $\mathbf{u}$) has an influence on the resulting scheme constructed with the above procedure.

In other words, to any permutation $p \in \mathcal{S}_N$, where $\mathcal{S}_N$ is the group of permutations of $\{1, \cdots, N\}$, we can associate a scheme of the same nature than (115) namely

$$\left| \begin{array}{l} \displaystyle \oint_\Omega^h \frac{u_\ell^{n+1} - 2u_\ell^n + u_\ell^{n-1}}{\Delta t^2} \cdot v_\ell + \oint_\Omega^h \delta_\ell H(\partial_x u_\ell^{n+1}, \partial_x u_\ell^{n-1}; \partial_x u_{j\neq\ell}^{n+sg(p(\ell)-p(j))}) \cdot \partial_x v_\ell = 0, \\ \hfill \forall\, v_\ell \in \mathcal{V}_\ell, \quad \ell = 1, \cdots, N. \end{array} \right. \tag{115}$$

which corresponds to the discrete gradient

$$\begin{array}{c} \mathbb{V}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \mathbb{V}^{(p)}(\mathbf{u}, \mathbf{v}, \mathbf{w}) \\ \left[\mathbb{V}^{(p)}(\mathbf{u}, \mathbf{v}, \mathbf{w})\right]_\ell := \left(\delta_\ell H(u_\ell, w_\ell; \left[\beta_{j\ell}^{(p)} u_j + (1 - \beta_{j\ell}^{(p)}) w_j\right]_{j\neq\ell}) \right) \end{array} \tag{116}$$

where by definition

$$\beta_{j\ell}^{(p)} = 1 \text{ if } p(j) < p(\ell), \quad 0 \text{ if } p(j) > p(\ell). \tag{117}$$

Any of these scheme will preserve the same energy (75) with $\mathbb{H}$ given by (114).

As an illustration, for $N = 2$, we obtain a scheme that differs from (111) by exchanging the roles of the indices 1 and 2:

$$\left\{ \begin{array}{l} \displaystyle \oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_1^{n-1}}{\Delta t^2} \cdot v_1 + \oint_\Omega^h \frac{H(\partial_x u_1^{n+1}, \partial_x u_2^{n+1}) - H(\partial_x u_1^{n-1}, \partial_x u_2^{n+1})}{\partial_x u_1^{n+1} - \partial_x u_1^{n-1}} \cdot \partial_x v_1 = 0, \\ \hfill \forall\, v_1 \in \mathcal{V}_1, \\ \displaystyle \oint_\Omega^h \frac{u_2^{n+1} - 2u_2^n + u_2^{n-1}}{\Delta t^2} \cdot v_2 + \oint_\Omega^h \frac{H(\partial_x u_1^{n-1}, \partial_x u_2^{n+1}) - H(\partial_x u_1^{n-1}, \partial_x u_2^{n-1})}{\partial_x u_2^{n+1} - \partial_x u_2^{n-1}} \cdot \partial_x v_2 = 0, \\ \hfill \forall\, v_2 \in \mathcal{V}_2. \end{array} \right. \tag{118}$$

A major drawback of these schemes is that each of them is not centered in time (the property (63) is not satisfied) and consequently only first order accurate in time. From this point of view, they can not be considered as a generalization of the θ-scheme.

To restore the second order accuracy that is valid for the scalar case, the idea is to take the "average" of these schemes, by choosing the average approximate gradient. For instance, in the case $N = 2$, the "average" of the schemes (111) and (118) gives

$$\left\{ \begin{array}{l} \displaystyle \oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_1^{n-1}}{\Delta t^2} \cdot v_1 + \frac{1}{2} \oint_\Omega^h \frac{H(\partial_x u_1^{n+1}, \partial_x u_2^{n+1}) - H(\partial_x u_1^{n-1}, \partial_x u_2^{n+1})}{\partial_x u_1^{n+1} - \partial_x u_1^{n-1}} \cdot \partial_x v_1 \\ \displaystyle \qquad + \frac{1}{2} \oint_\Omega^h \frac{H(\partial_x u_1^{n+1}, \partial_x u_2^{n-1}) - H(\partial_x u_1^{n-1}, \partial_x u_2^{n-1})}{\partial_x u_1^{n+1} - \partial_x u_1^{n-1}} \cdot \partial_x v_1 = 0, \\ \displaystyle \oint_\Omega^h \frac{u_2^{n+1} - 2u_2^n + u_2^{n-1}}{\Delta t^2} \cdot v_2 + \frac{1}{2} \oint_\Omega^h \frac{H(\partial_x u_1^{n+1}, \partial_x u_2^{n+1}) - H(\partial_x u_1^{n+1}, \partial_x u_2^{n-1})}{\partial_x u_2^{n+1} - \partial_x u_2^{n-1}} \cdot \partial_x v_2 \\ \displaystyle \qquad + \frac{1}{2} \oint_\Omega^h \frac{H(\partial_x u_1^{n-1}, \partial_x u_2^{n+1}) - H(\partial_x u_1^{n-1}, \partial_x u_2^{n-1})}{\partial_x u_2^{n+1} - \partial_x u_2^{n-1}} \cdot \partial_x v_2 = 0. \end{array} \right. \tag{119}$$

The generalization to any N consists in taking

$$\nabla H(\mathbf{u},\mathbf{v},\mathbf{w}) = \frac{1}{N!}\sum_{p\in\mathcal{S}_N} \nabla H^{(p)}(\mathbf{u},\mathbf{v},\mathbf{w}). \tag{120}$$

Clearly, this discrete gradient is conservative still with the discrete energy given by (114). Moreover, it is centered in time (and then second order accurate) because (63) is satisfied. To check (63) easily we introduce the bijection $\mathcal{I}$ from $\mathcal{S}_N$ into itself

$$p \mapsto q = \mathcal{I}(p) \quad \text{such that} \quad q(j) = p(N+1-j) \tag{121}$$

Then, according to (120) and (116), we write:

$$\Big[\nabla H(\mathbf{w},\mathbf{v},\mathbf{u})\Big]_\ell = \frac{1}{N!}\sum_{p\in\mathcal{S}_N} \big(\, \delta_\ell H(w_\ell,u_\ell;\big[\,\beta^{(p)}_{j\ell} w_j + (1-\beta^{(p)}_{j\ell})u_j\,\big]_{j\neq\ell})\,\big)$$

We write $p=\mathcal{I}(q)$ in the sum (so that q describes $\mathcal{S}_N$ when p describes $\mathcal{S}_N$) and notice that $p=\mathcal{I}(q) \Longrightarrow \beta^{(p)}_{j\ell} = 1-\beta^{(q)}_{j\ell}$ to write

$$\Big[\nabla H(\mathbf{w},\mathbf{v},\mathbf{u})\Big]_\ell = \frac{1}{N!}\sum_{q\in\mathcal{S}_N} \big(\, \delta_\ell H(w_\ell,u_\ell;\big[\, + (1-\beta^{(q)}_{j\ell}) w_j + \beta^{(p)}_{j\ell}u_j\,\big]_{j\neq\ell})\,\big)$$

Since $\delta_\ell H(u_\ell,w_\ell;v_{j\neq\ell})$ is symmetric in (u_ℓ,w_ℓ) we have

$$\Big[\nabla H(\mathbf{w},\mathbf{v},\mathbf{u})\Big]_\ell = \frac{1}{N!}\sum_{q\in\mathcal{S}_N} \big(\, \delta_\ell H(u_\ell,w_\ell;\big[\,\beta^{(q)}_{j\ell} u_j + (1-\beta^{(q)}_{j\ell})w_j + \,\big]_{j\neq\ell})\,\big) = \Big[\nabla H(\mathbf{u},\mathbf{v},\mathbf{w})\Big]_\ell.$$

In formula (120), appears as a sum of $N!$ terms. However, we are going to see that, componemt by component, it can be rewritten it as the sum of only 2^{N_1} terms, which will be useful for the practical implementation of the scheme.

To state the result, it is useful to introduce the sets:

$$J_\ell = \{1,\dots,N\}\setminus\{\ell\}, \quad \ell = 1,\cdots,N,$$

and for each $1\leq\ell\leq N$

$$\Sigma_\ell = \{\sigma : J_\ell \longrightarrow \{+1,-1\}\,\},$$

the set of applications from J_ℓ into $\{+1,-1\}$ (that contains 2^{N-1} elements). Finally to each $\sigma\in\Sigma_\ell$, we associate the integer $\mu(\sigma)$ defined by

$$\mu(\sigma) = \#\,\big\{l\in J_k,\, \sigma(l)=+1\big\} = \#\sigma^{-1}(+1).$$

Lemma 2.2.6 *The approximate gradient defined by (120) is also given by*

$$\Big[\nabla H(\mathbf{u},\mathbf{v},\mathbf{w})\Big]_\ell = \sum_{\sigma\in\Sigma_\ell} \zeta(\sigma)\;\delta_\ell H\Big(u_\ell.w_\ell;\Big[\,(\frac{1+\sigma(j)}{2})\,u_j + (\frac{1-\sigma(j)}{2})\,w_j\,\Big]_{j\neq\ell}\Big) \tag{122}$$

where $\zeta(\sigma) = \dfrac{\mu(\sigma)!\,\big(N-1-\mu(\sigma)\big)!}{N!}$.

Proof. Let us introduce the map

$$\begin{array}{rcccc} \Phi_\ell & : & \mathcal{S}_N & \longrightarrow & \Sigma_\ell \\ & & p & \mapsto & \Phi_\ell(p)=\sigma^\ell_p \end{array} \tag{123}$$

where σ_p^ℓ is defined by

$$\forall\, j \in \Sigma_\ell, \quad \sigma_p^\ell(j) = sg\big(p(\ell) - p(j)\big) \tag{124}$$

We can rearrange the sum (120) as:

$$\Big[\,\nabla H(\mathbf{u},\mathbf{v},\mathbf{w})\,\Big]_\ell = \frac{1}{N!}\sum_{\sigma\in\mathcal{S}_\ell}\ \sum_{p\in\Phi_\ell^{-1}(\sigma)} \Big[\,\nabla H^{(p)}(\mathbf{u},\mathbf{v},\mathbf{w})\,\Big]_\ell.$$

Next, we remark that, inside each level set of Φ_ℓ, namely $\Phi_\ell^{-1}(\sigma)$, $\Big[\,\nabla H^{(p)}(\mathbf{u},\mathbf{v},\mathbf{w})\,\Big]_\ell$ is independent of p. Indeed, from the definition (117) we see that

$$\beta_{j\ell}^{(p)} = \Big(1+\sigma_p^\ell(j)\Big)/2.$$

Therefore, we deduce that ($\sigma_p^\ell = \sigma$ when p describes $\Phi_\ell^{-1}(\sigma)$)

$$\forall\, p \in \Phi_\ell^{-1}(\sigma), \quad \Big[\,\nabla H^{(p)}(\mathbf{u},\mathbf{v},\mathbf{w})\,\Big]_\ell = \delta_\ell H\Big(u_\ell, w_\ell; \Big[\,(\frac{1+\sigma(j)}{2})\,u_j + (\frac{1-\sigma(j)}{2})\,w_j\,\Big]_{j\neq\ell}\Big)$$

which yields

$$\Big[\nabla H(\mathbf{u},\mathbf{v},\mathbf{w})\Big]_\ell = \sum_{\sigma\in\mathcal{S}_\ell}\Big[\#\Phi_\ell^{-1}(\sigma)\Big]\ \delta_\ell H\Big(u_\ell, w_\ell; \Big[\,(\frac{1+\sigma(j)}{2})\,u_j + (\frac{1-\sigma(j)}{2})\,w_j\,\Big]_{j\neq\ell}\Big).$$

To conclude it suffices to show that

$$\#\Phi_\ell^{-1}(\sigma) = \mu(\sigma)!\big(N-1-\mu(\sigma)\big)!. \tag{125}$$

To check this, given $\sigma \in \Sigma_\ell$, we set $m = \mu(\sigma) \in \{\,1,\,\dots,N-1\}$ and

$$\mathcal{I}_+ = \{j \in J_\ell \,|\, \sigma(j) = +1\}\ (\#\mathcal{I}_+ = m) \quad \text{and} \quad \mathcal{I}_- = \{j \in J_\ell \,|\, \sigma(j) = -1\}\ (\#\mathcal{I}_- = N-1-m).$$

Next it suffices to remark that

$$(i)\ \Phi_\ell(p) = \sigma \iff (ii)\ \begin{cases} p(\ell) = m+1 \\ p|_{\mathcal{I}_+} \text{ is a bijection from } \mathcal{I}_+ \text{ into } \{\,1,\cdots,m\,\} \\ p|_{\mathcal{I}_-} \text{ is a bijection from } \mathcal{I}_- \text{ into } \{\,m+2,\cdots,N\,\} \end{cases} \tag{126}$$

Indeed, $\Phi_\ell(p) = \sigma$ means that when j describes $\mathcal{I}_+$ (resp. j describes $\mathcal{I}_-$), $p(j)$ takes m values strictly smaller than $p(\ell)$ (resp. $N-1-m$ values strictly greater than $p(\ell)$). As a consequence, the only possibility is $p(\mathcal{I}_+) = \{\,1,\cdots,m\,\}$, $p(\ell) = m+1$ and $p(\mathcal{I}_+) = \{\,m+2,\cdots,N\,\}$. This proves $(i) \Rightarrow (ii)$. The recerve statement is obvious.

To count the number of antecedents of σ via Φ_ℓ, from (126), it suffices to multiply the numbers of bijections in a set with m elements with the numbers of bijections in a set with $N-1-m$ elements, which leads to (125).

□

Finally, the equations of scheme associated to (120) (or (122)) are:

$$\boxed{\oint_\Omega^h \frac{u_\ell^{n+1} - 2u_\ell^n + u_\ell^{n-1}}{\Delta t^2}\cdot v_\ell + \sum_{\sigma\in\Sigma_\ell}\zeta(\sigma)\oint_\Omega^h \delta_\ell H(\partial_x u_\ell^{n+1}, \partial_x u_\ell^{n-1}; \partial_x u_{j\neq\ell}^{n+\sigma(\ell)})\cdot\partial_x v_\ell = 0, \quad \forall\, v_\ell \in \mathcal{V}_\ell, \quad \ell = 1,\,\dots,\,N.} \tag{127}$$

To give a concrete example of this scheme, it leads in the particular case $N = 3$ to the following scheme:

$$\left\{\begin{aligned}
&\text{Find } (u_1, u_2, u_3) \in \mathcal{V}_h \text{ such that for any } (v_1, v_2, v_3) \in \mathcal{V}_h,\\
&\oint_\Omega^h \frac{u_1^{n+1} - 2u_1^n + u_1^{n-1}}{\Delta t^2}\, v_1 + \frac{1}{6}\oint_\Omega^h \Big[2\,\delta_1 H(\partial_x u_1^{n+1}, \partial_x u_1^{n-1}; \partial_x u_2^{n+1}, \partial_x u_3^{n+1}) + \\
&\qquad 2\,\delta_1 H(\partial_x u_1^{n+1}, \partial_x u_1^{n-1}; \partial_x u_2^{n-1}, \partial_x u_3^{n-1}) + \\
&\qquad \delta_1 H(\partial_x u_1^{n+1}, \partial_x u_1^{n-1}; \partial_x u_2^{n+1}, \partial_x u_3^{n-1}) + \\
&\qquad \delta_1 H(\partial_x u_1^{n+1}, \partial_x u_1^{n-1}; \partial_x u_2^{n-1}, \partial_x u_3^{n+1}) \Big]\, \partial_x v_1 = 0,\\
&\oint_\Omega^h \frac{u_2^{n+1} - 2u_2^n + u_2^{n-1}}{\Delta t^2}\, v_2 + \frac{1}{6}\oint_\Omega^h \Big[2\,\delta_2 H(\partial_x u_2^{n+1}, \partial_x u_2^{n-1}; \partial_x u_1^{n+1}, \partial_x u_3^{n+1}) + \\
&\qquad 2\,\delta_2 H(\partial_x u_2^{n+1}, \partial_x u_2^{n-1}; \partial_x u_1^{n-1}, \partial_x u_3^{n-1}) + \\
&\qquad \delta_2 H(\partial_x u_2^{n+1}, \partial_x u_2^{n-1}; \partial_x u_1^{n+1}, \partial_x u_3^{n-1}) + \\
&\qquad \delta_2 H(\partial_x u_2^{n+1}, \partial_x u_2^{n-1}; \partial_x u_1^{n-1}, \partial_x u_3^{n+1}) \Big]\, \partial_x v_2 = 0,\\
&\oint_\Omega^h \frac{u_3^{n+1} - 2u_3^n + u_3^{n-1}}{\Delta t^2}\, v_3 + \frac{1}{6}\oint_\Omega^h \Big[2\delta_3 H(\partial_x u_3^{n+1}, \partial_x u_3^{n-1}; \partial_x u_1^{n+1}, \partial_x u_2^{n+1}) + \\
&\qquad 2\,\delta_3 H(\partial_x u_3^{n+1}, \partial_x u_3^{n-1}; \partial_x u_1^{n-1}, \partial_x u_2^{n-1}) + \\
&\qquad \delta_3 H(\partial_x u_3^{n+1}, \partial_x u_3^{n-1}; \partial_x u_1^{n+1}, \partial_x u_2^{n-1}) + \\
&\qquad \delta_3 H(\partial_x u_3^{n+1}, \partial_x u_3^{n-1}; \partial_x u_1^{n-1}, \partial_x u_2^{n+1}) \Big]\, \partial_x v_3 = 0.
\end{aligned}\right.$$

This paragraph tackles the issue of time discretization, in order to achieve numerical stability by energy preservation. The main idea is to generalize the θ-schemes. We focused on three-point schemes, and tried to find a discrete expression of the gradient of the potential energy, which is consistent and allows an energy preservation.

We introduced a definition of "conservative" approximate gradient function and scheme (definition 2.1) that leads to a simple energy preservation (lemma 2.2.1), and unconditional L^2 stability of the discrete solution (corollary 76). We found out (lemma 2.2.2) that the explicit scheme leads to a conservative scheme if and only if the system of equations is linear.

In order to build preserving schemes for all size of system, we first handled scalar equations, and introduced a particular function (96) on which are built preserving schemes (97) and (98). We then tried to generalize this idea to greater size of systems. The most naïve idea (101) (partially implicit scheme) has been shown to be preserving only for very particular potential energies (lemma 2.2.4). A triangular implicit scheme (115) has been shown to be preserving of a simple energy (lemma 2.2.5), but only first order accurate. A generalization which is second order accurate has been studied (120) and we found a simple expression (127) for this general preserving second order accurate scheme.

3 Numerical results for the nonlinear string

3.1 Numerical issues

3.1.1 The computational algorithm

Finite element discretization Concretely, in the case of the nonlinear string, Ω is a segment, noted $[0, L]$. For each direction, we discretize $H_0^1\big([0, L]\big)$ with P_k elements : dividing the domain into $N_x - 1$ elements, each element has $k + 1$ secondary points constituting the lattice. Basis functions are continuous, piecewise polynomials of degree k (ie polynomial on each element), being 1 at one point and 0 at every others. We can then write the system, equivalent to the discrete variational formulation (42), taking as test functions in (127) certain basis functions of $\mathcal{X}_h$, chosen in order to achieve equivalency.

We can evaluate the dimension of $\mathcal{X}_h$, that is to say the number of degrees of freedom, which depends on the degree k of the chosen polynoms, on the number of elements $N_x - 1$ in the mesh, and the size N of the system. The account leads to

$$N_h = N \times N_d = N[(N_x - 1)k + 1]$$

degrees of freedom, which leads to the same number of lines in the corresponding system. This system is nonlinear, which prevents us to write the problem in the usual matricial form.

Nonlinear resolution Programming the scheme (127) amounts to nullify, at each time step, a function $F : \mathbb{R}^{N_h} \longrightarrow \mathbb{R}^{N_h}$ a priori highly non linear. The dimension N_h depends on the size of the continuous system, the number of finite elements taken for the resolution, but also on the order of the chosen polynoms. The method we use to find a zero of F is Newton's method, which consists in going from an initial point, then invert the jacobian matrix of F until we find a point $\mathbf{U}^*$ such that $F(\mathbf{U}^*)$ is "close enough" to zero. We have to calculate this jacobian matrix, which depends on the point at which we estimate it, since the problem is nonlinear.

We can write the scheme's solution $\mathbf{u}_h$ as its decomposition on basis functions:

$$\mathbf{u}_h = \sum_{\ell=1}^{N} \sum_{j=1}^{N_d} u_{\ell,j} \psi_{\ell,j}$$

where $\psi_{\ell,j}$ is a vector having only one non zero component, directed along the direction ℓ, which is ϕ_j ie the basis function P_k associated with the degree of freedom j. There are N_d degrees of freedom for each direction. The unknown of the problem, at each time step, is the vector $\mathbf{U}_h = \big(u_{\ell,j}^{n+1}\big)_{\ell,j}$, the values $\big(u_{\ell,j}^{n}\big)_{\ell,j}$ and $\big(u_{\ell,j}^{n-1}\big)_{\ell,j}$ being considered known and constant. Then, the scheme consists in N_h lines, and the line corresponding to the direction ℓ and the degree of freedom j can be written:

$$F_{\ell,j}(\mathbf{U}_h) = \oint_{[0\,L]}^{h} \frac{u_\ell^{n+1} - 2u_\ell^n + u_\ell^{n-1}}{\Delta t^2} \phi_j + \oint_{[0\,L]}^{h} \sum_{\sigma \in \Sigma_\ell} \zeta(\sigma)\, \delta_\ell H(\partial_x u_\ell^{n+1}, \partial_x u_\ell^{n-1}; \partial_x u_{\tilde{\ell} \neq \ell}^{n+\sigma(\tilde{\ell})}) \partial_x \phi_j$$

where $u_\ell = \displaystyle\sum_{p=1}^{N_d} u_{\ell,p} \phi_p$.

The jacobian of this scheme is then a matrix of the applications $\dfrac{\partial F_{\ell,j}}{\partial u_{k,n}}$ from $\mathbb{R}^{N_h}$ to $\mathbb{R}$:

$$\begin{cases} \dfrac{\partial F_{\ell,j}}{\partial u_{\ell,n}} &= \dfrac{1}{\Delta t^2}\oint_0^L \phi_n\phi_j + \displaystyle\sum_{\sigma\in\Sigma_\ell}\zeta(\sigma)\int_0^L \dfrac{\partial \delta_\ell H}{\partial \ell}(\partial_x u_\ell^{n+1}, \partial_x u_\ell^{n-1}; \partial_x u_{\tilde{\ell}\neq\ell}^{n+\sigma(\tilde{\ell})})\partial_x\phi_n\partial_x\phi_j \\ \dfrac{\partial F_{\ell,j}}{\partial u_{k,n}} &= \displaystyle\sum_{\substack{\sigma\in\Sigma_\ell\\ \sigma(k)=+1}}\zeta(\sigma)\int_0^L \dfrac{\partial \delta_\ell H}{\partial k}(\partial_x u_\ell^{n+1}, \partial_x u_\ell^{n-1}; \partial_x u_{\tilde{\ell}\neq\ell}^{n+\sigma(\tilde{\ell})})\partial_x\phi_n\partial_x\phi_j \end{cases}$$

These quantities clearly depends on the point where we calculate and not only, as in the linear case, on the basis functions and time, space steps choice.

3.1.2 Nonlinear string model : specific difficulties

Calculation of $\delta_\ell H$ The calculation of the functions $\delta_\ell H$ is quite tricky on a numerical point of view since their definition is:

$$\delta_\ell H(u_\ell, \tilde{u}_\ell; u_{\tilde{\ell}\neq\ell}) = \begin{cases} \dfrac{H(u_\ell, u_{\tilde{\ell}\neq\ell}) - H(\tilde{u}_\ell, u_{\tilde{\ell}\neq\ell})}{u_\ell - \tilde{u}_\ell} & \text{if } u_\ell \neq \tilde{u}_\ell, \\ \dfrac{\partial H}{\partial \ell}(u_1, \ldots, u_\ell, \ldots, u_N) & \text{if } u_\ell = \tilde{u}_\ell. \end{cases} \tag{128}$$

Indeed, numerically, the calculation of a difference between two numbers is exact until the two numbers are closer than the numerical precision. Under this precision, calculation will give numerical instabilities. Consequently, two numbers are considered equal, in the code, when their relative difference is lower than a chosen tolerance, which artificially creates a gap, a discontinuity in the values taken by the function $\delta_\ell H$. Some manipulations have shown that when the analytic expression of H is known and nice, it is sometimes possible write the function $\delta_\ell H$ in a way that eludes the singularity of the denominator when u_k and $\tilde{u}_k$ become close (with no division), which is precisely the cause of the numerical problem. This remark is true in the particular case of H_{ex}, for $N = 2$ and $N = 3$ as well. Let $\mathbf{u} = (u_1, \ldots, u_{N-1}, v)$, and

$$H_{ex}(\mathbf{u}) = \frac{1}{2}\sum_{k=1}^{N-1} u_k^2 + \frac{1}{2}v^2 - \alpha\Big[\sqrt{\sum_{k=1}^{N-1} u_k^2 + (1+v)^2} - (1+v)\Big]$$

We can rewrite the directional finite derivative functions with no division by $u_k - \tilde{u}_k$, as:

$$\begin{cases} \delta_\ell H_{ex}(u_\ell, \tilde{u}_\ell; u_{l\neq\ell}) = \dfrac{1}{2}(u_\ell + \tilde{u}_\ell) - \alpha\dfrac{u_\ell + \tilde{u}_\ell}{\sqrt{\sum_{k=1}^{N-1} u_k^2 + (1+v)^2} + \sqrt{\sum_{k=1}^{N-1} \tilde{u}_k^2 + (1+v)^2}} & \forall\, \ell \neq N, \\ \delta_N H_{ex}(v, \tilde{v}; u_{l\neq N}) = \dfrac{1}{2}(v + \tilde{v}) + \alpha - \alpha\dfrac{2 + v + \tilde{v}}{\sqrt{\sum_{k=1}^{N-1} u_k^2 + (1+v)^2} + \sqrt{\sum_{k=1}^{N-1} u_k^2 + (1+\tilde{v})^2}}. \end{cases}$$

These expressions tend towards the directional partial derivative when $\tilde{u}_k$ tends towards u_k, which avoids to program multiple cases for the corresponding expressions. We can also express the derivatives of these functions, which we will need for the jacobian calculation.

$$\left\{\begin{array}{lcll}
\partial_j \delta_\ell H_{ex}(u_\ell,\tilde{u}_\ell;u_{l\neq\ell}) & = & \alpha\,\dfrac{(u_\ell+\tilde{u}_\ell)u_j\left[\frac{1}{\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}}+\frac{1}{\sqrt{\sum_{k=1}^{N-1}\tilde{u}_k^2+(1+v)^2}}\right]}{\left[\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}+\sqrt{\sum_{k=1}^{N-1}\tilde{u}_k^2+(1+v)^2}\ \right]^2} & \forall\ k \text{ and } j\neq N,\\
\partial_\ell \delta_\ell H_{ex}(u_\ell,\tilde{u}_\ell;u_{l\neq\ell}) & = & \frac{1}{2}-\alpha\left[\frac{1}{\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}+\sqrt{\sum_{k=1}^{N-1}\tilde{u}_k^2+(1+v)^2}}-\right. & \\
& & \left.\dfrac{(u_\ell+\tilde{u}_\ell)u_\ell}{\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}\left[\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}+\sqrt{\sum_{k=1}^{N-1}\tilde{u}_k^2+(1+v)^2}\ \right]^2}\right] & \forall\ k\neq N,\\
\partial_N \delta_\ell H_{ex}(u_\ell,\tilde{u}_\ell;u_{l\neq\ell}) & = & \alpha\,\dfrac{(u_\ell+\tilde{u}_\ell)(1+v)\left[\frac{1}{\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}}+\frac{1}{\sqrt{\sum_{k=1}^{N-1}\tilde{u}_k^2+(1+v)^2}}\right]}{\left[\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}+\sqrt{\sum_{k=1}^{N-1}\tilde{u}_k^2+(1+v)^2}\ \right]^2} & \forall\ k\neq N,\\
\partial_j \delta_N H_{ex}(v,\tilde{v};u_{l\neq N}) & = & \alpha\,\dfrac{(2+v+\tilde{v})u_j\left[\frac{1}{\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}}+\frac{1}{\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+\tilde{v})^2}}\right]}{\left[\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}+\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+\tilde{v})^2}\ \right]^2} & \forall\ k\neq N\\
\partial_N \delta_N H_{ex}(v,\tilde{v};u_{l\neq N}) & = & \frac{1}{2}-\alpha\left[\frac{1}{\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}+\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+\tilde{v})^2}}-\right. & \\
& & \left.\dfrac{(2+v+\tilde{v})(1+v)}{\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}\left[\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+v)^2}+\sqrt{\sum_{k=1}^{N-1}u_k^2+(1+\tilde{v})^2}\ \right]^2}\right] &
\end{array}\right.$$

The expected following properties, which are true with the definition (128), are still satisfied with these (equivalent) definitions:

$$\left\{\begin{array}{llll}
\partial_j \delta_\ell H_{ex}(u_\ell,\tilde{u}_\ell;u_{l\neq\ell}) & \xrightarrow{\tilde{u}_\ell\to u_\ell} & \partial^2_{j,\ell}H_{ex} & \forall\ k \text{ and } j\neq N,\\
\partial_\ell \delta_\ell H_{ex}(u_\ell,\tilde{u}_\ell;u_{l\neq\ell}) & \xrightarrow{\tilde{u}_\ell\to u_\ell} & \dfrac{1}{2}\,\partial^2_{\ell,\ell}H_{ex} & \forall\ \ell\neq N,\\
\partial_N \delta_\ell H_{ex}(u_\ell,\tilde{u}_\ell;u_{l\neq\ell}) & \xrightarrow{\tilde{u}_\ell\to u_\ell} & \partial^2_{N,\ell}H_{ex} & \forall\ \ell\neq N,\\
\partial_j \delta_N H_{ex}(v,\tilde{v};u_{l\neq N}) & \xrightarrow{\tilde{v}\to v} & \partial^2_{j,N}H_{ex} & \forall\ k\neq N\\
\partial_N \delta_N H_{ex}(v,\tilde{v};u_{l\neq N}) & \xrightarrow{\tilde{v}\to v} & \dfrac{1}{2}\,\partial^2_{N,N}H_{ex} &
\end{array}\right.$$

This paragraph laid the emphasis on numerical choices that we made : we chose standard approximation with Lagrange finite elements of degree k, which leads to a numerical problem counting N_h unknowns. The scheme amounts to nullify a nonlinear function $F:\mathbb{R}^{N_h}\to\mathbb{R}^{N_h}$, we chose to use a Newton's method to approximate a zero. Finally, the nonlinear string model based on the geometrically exact model allows us to rewrite the functions $\delta_\ell H$, defined with a formal singularity by (128), as regular functions on the numerical point of view.

3.2 Tests of numerical results for the string model

We have implemented the nonlinear string model described in section 1.1 (geometrically exact model) using our preserving scheme (127), and the developed Bank-Sujbert model. The numerical results were interesting to show the influence of nonlinearity on the behavior of a vibrating string, and to show the comparison between exact and developed models.

3.2.1 Influence of the nonlinearity

Numerical experiments have been made on very simple problems to show the influence of the nonlinearity on the string vibration. Nonlinear behavior can come from two different factors :

either the nonlinear factor α comes closer to 1, either the initial amplitude grows. Figure (4) shows the deformation of a string with a sinus as initial condition on the transversal direction. The time is represented as the third space variable, and the different snapshots are displayed in different colors from blue to red. The nonlinear factor is changed between the lines, and the initial amplitude of the sinus is changed between the columns.

The first line has been made with a nonlinear factor $\alpha = 0$, which leads to two linear uncoupled wave equations. Indeed, the initial amplitude of the data has just a scaling influence on the vibration of the string, expressing the linear behavior of the solution.

The first column shows, for small initial data, the influence of the nonlinear factor α. We can notice that the vibration is much slower while α increases, which is in agreement with the second order Taylor expansion of the potential energy H_{ex}: we indeed get as approximated system two uncoupled linear wave equations, with a celerity of 1 for the longitudinal wave and $\sqrt{1-\alpha}$ for the transversal wave, which is the one that we can observe. For $\alpha = 0$ the celerity is $c = 1$ and for $\alpha = 0.99$ the celerity is $c = 0.1$, ten times less.

Finally, if we look at the last two lines, we can see that increasing the initial amplitude leads to non usual behaviors of the string, pointing out the nonlinear influence of the equation, and especially the stretching of the string due to the presence of longitudinal waves.

We can add that the simulations presented here have lead to a very good energy preservation (about 10^{-13} relative error on the energy preservation for a Newton tolerance of 10^{-13} on the ℓ^2 norm of $F(\mathbf{U}_h)$).

3.2.2 Comparison with approximate Bank-Sujbert model

Another interesting point was to compare the string deformation when we use the geometrically exact model and when we use the Taylor expansion used in [2] and [4]. Our scheme makes it easy to switch from one model to another, and the result of simulation can be seen in figure (5).

The system of lines and columns is exactly the same as the previous figure : different nonlinear factors α in lines, different initial amplitudes in columns. Each subfigure shows the string deformation by exact model in red (solid line) and in expanded model in blue (dashed line).

For $\alpha = 0$ (the first line) the two models coincide, and we can see that the simulations give the same result (red (solid) curve and blue (dashed) curve are the same).

But for more realistic values of α (for real piano strings it can be more than 0.999) the two curves are about the same for a very small initial amplitude, but are slightly different for $|u_0| = 0.1$ and very different for $|u_0| = 0.3$.

These two numerical examples show the influence of the nonlinearity on the behavior of the vibrating string, and a comparison between exact and developed models.

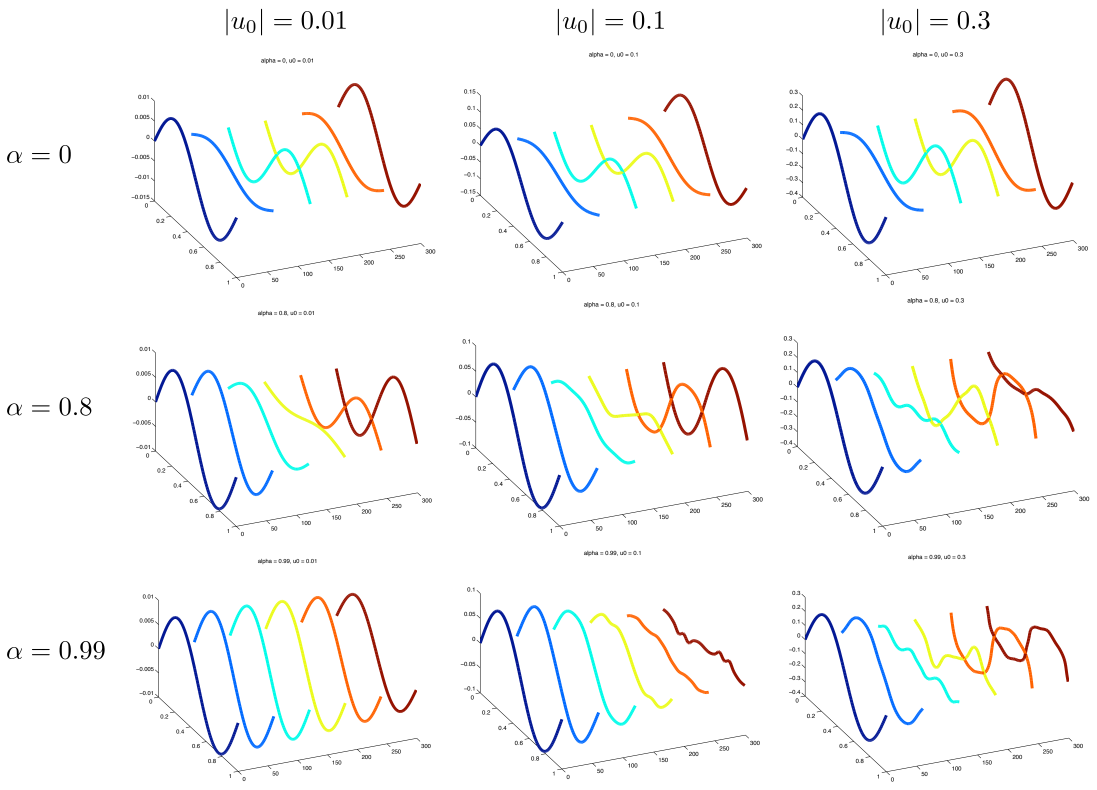


Figure 4: Evolution of the string deformation (time evolution in each subfigure from blue (left) to red (right)), for different values of α and initial amplitude.

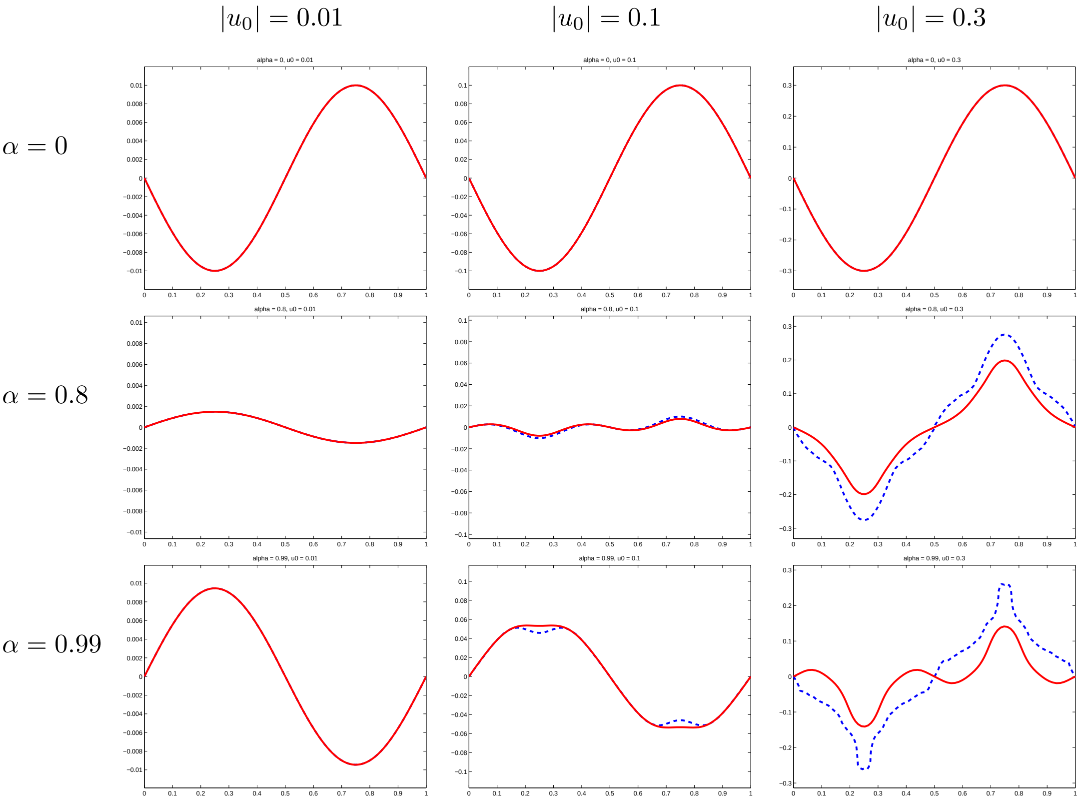


Figure 5: Comparison of a snapshot of the string deformation for exact (red, solid) and approximate (blue, dashed) model, for different values of α and initial amplitude.

Conclusions and perspectives

The problem was to model the piano string with nonlinear coupling between transversal and longitudinal directions of vibration. We derived the nonlinear geometrically exact model (GEM), and we have seen that any stress-strain law, not necessarily affine as Hooke's, leads to the same structure called "hamiltonian system of wave equations". Numerous approximate models (including Bank-Sujbert model) come from the Taylor expansions of the potential energy of the GEM.

We introduced a class of systems differing from each other in the expression of their potential energy H, in which fitted the GEM (associated with H_{ex}) and its approximation including Bank-Sujbert model (associated with H_{BS}). We have shown some properties inherent to this class of systems, assuming hypothesis on the model. The main properties are energy preservation, H^1 stability of the solution if H is greater than a parabola (coercivity property), local hyperbolicity if and only if H is locally convex, existence and unicity of classical smooth solution for small initial data if the hessian matrix D^2H is linearly degenerated, finite propagation velocity and symmetry preservation. Only the GEM satisfied all of them.

On a numerical point of view, we have tried to find schemes that preserve a discrete energy, in a physical purpose but also in order to guarantee numerical stability, which is not an easy task for nonlinear problems. The semi discretization in space that we choose is classic, based on a variational formulation and the choice of a family of finite dimensional subspaces. This semi discretization preserves a semi discrete energy by construction. Time discretization is the difficult point when we try to build preserving schemes for any size of system. The approach we lead was guided by the case of scalar equation. The explicit scheme has been rejected since it cannot be preserving unless the continuous problem is linear. A partially implicit scheme (implicit only for the variable of the considered line in the system) has been shown to be preserving only for very particular potential energies. A triangular implicit scheme has been introduced, preserving for any potential energy, but only first order accurate. Finally, a fully implicit was proposed, which preserves a simple discrete potential energy and is second order accurate.

Some numerical results were exhibited for the nonlinear string vibration. A new challenge is that stiff string does not fit the hamiltonian system of wave equations class that we have considered, we will in a near future develop similar preserving schemes for such problems. In the objective of modeling a whole piano, we also need to derive a hammer model, a bridge model, a soundboard and create a preserving numerical scheme for the whole structure coupled with air. The energy approach should enable us to achieve stability for the coupled problem without digging further.

A Why the naïve scheme is very restrictive for $N \geq 3$

In order to preserve a discrete energy, we want to build a function $\mathbb{H} : (x, y) \in \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ and a function $\mathbb{\nabla} H : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$ such that:

$$\frac{\mathbb{H}(\mathbf{u}^{n+1}, \mathbf{u}^n) - \mathbb{H}(\mathbf{u}^n, \mathbf{u}^{n-1})}{\Delta t} = \mathbb{\nabla} H(\mathbf{u}^{n+1}, \mathbf{u}^n, \mathbf{u}^{n-1}) \cdot \frac{\mathbf{u}^{n+1} - \mathbf{u}^{n+1}}{2\Delta t} \tag{129}$$

where $\mathbb{H}$ is consistent with the continuous energy H, ie : $\mathbb{H}(\mathbf{u}, \mathbf{u}) = H(\mathbf{u})$, and $\mathbb{\nabla} H$ is consistent with the continuous gradient ∇H. A natural way of trying to solve this problem is to handle the k^{th} line of the gradient $\mathbb{\nabla} H$ as the directional finite difference in the variable k, letting all the other variables at time n:

$$[\mathbb{\nabla} H]_k \left(\mathbf{u}^{n-1}, \mathbf{u}^n, \mathbf{u}^{n+1}\right) = \frac{H(u_k^{n+1}; u_{l\neq k}^n) - H(u_k^{n-1}; u_{l\neq k}^n)}{u_k^{n+1} - u_k^{n+1}} \tag{130}$$

We have called this proposition the "naïve scheme" since it is a natural extension of the first idea for $N = 1$.

The question that we raise here is to know if there always exists a function $\mathbb{H}$ associated with this way of writing $\mathbb{\nabla} H$; and if the answer is yes, what form should have this function. The constraint on $\mathbb{H}$ comes from (130) and (129):

$$\boxed{\mathbb{H}(\mathbf{u}^{n+1}, \mathbf{u}^n) - \mathbb{H}(\mathbf{u}^n, \mathbf{u}^{n-1}) = \frac{1}{2} \sum_{1 \leq k \leq N} \left[H(u_k^{n+1}; u_{l\neq k}^n) - H(u_k^{n-1}; u_{l\neq k}^n)\right]} \tag{131}$$

A.1 Naïve scheme for scalar equation : a necessary form

In the scalar case, $N = 1$ and the problem (131) becomes : Find a function $\mathbb{H} : \mathbb{R}^2 \to \mathbb{R}$ such that:

$$\mathbb{H}(x, y) - \mathbb{H}(y, z) = \frac{H(x) - H(z)}{2} \qquad \forall\, (x, y, z) \in \mathbb{R}^3 \tag{132}$$

If such a function exists, we cake derive the expression (132) along x:

$$\partial_1 \mathbb{H}(x, y) = \frac{1}{2} H'(x) \Rightarrow \text{There exists a function } \phi \text{ and } c \in \mathbb{R} \mid \mathbb{H}(x, y) = \frac{1}{2} H(x) + \phi(y) + c$$

Knowing that if $\mathbb{H}$ exists, is must be under this form, we can try it in (132):

$$\left[\frac{1}{2} H(x) + \phi(y) + c\right] - \left[\frac{1}{2} H(y) + \phi(z) + c\right] = \frac{H(x) - H(z)}{2} \qquad \forall\, (x, y, z) \in \mathbb{R}^3 \tag{133}$$

Since the right member does not depend on y, we have : $\phi(y) = \frac{1}{2} H(y)$. Then

$$\boxed{\mathbb{H}(x, y) = \frac{1}{2}\left[H(x) + H(y)\right] + c} \tag{134}$$

It is easy to verify that (132) is satisfied.

The conclusion is that, if $N = 1$, for any function H you can find functions $\mathbb{H}$ such that (132) is satisfied. These functions $\mathbb{H}$ that satisfy (132) are written as (134).

A.2 Naïve scheme for a two lines system : a necessary form

The problem (131) becomes : Find a function $\mathbb{H} : \mathbb{R}^4 \to \mathbb{R}$ such that, $\forall\,(x_1, x_2, y_1, y_2, z_1, z_2) \in \mathbb{R}^6$,

$$\mathbb{H}(x_1, x_2; y_1, y_2) - \mathbb{H}(y_1, y_2; z_1, z_2) = \frac{H(x_1, y_2) - H(z_1, y_2) + H(y_1, x_2) - H(y_1, z_2)}{2} \tag{135}$$

Since this constraint must be true for all values of its arguments, we can first take $x_2 = y_2 = z_2 = \lambda$, and we obtain:

$$\mathbb{H}(x_1, \lambda; y_1, \lambda) - \mathbb{H}(y_1, \lambda; z_1, \lambda) = \frac{H(x_1, \lambda) - H(z_1, \lambda) + \cancel{H(y_1, \lambda)} - \cancel{H(y_1, \lambda)}}{2}$$

Some of the terms disappear, and we have to deal with a problem similar to (132). We can use the conclusion of the previous paragraph to establish that there exists a function $c_1(\lambda)$ such that:

$$\mathbb{H}(x_1, \lambda; y_1, \lambda) = \frac{H(x_1, \lambda) + H(y_1, \lambda)}{2} + c_1(\lambda) \tag{136}$$

Similarly, we can take $x_1 = y_1 = z_1 = \mu$ in (135) and we obtain:

$$\mathbb{H}(\mu, x_2; \mu, y_2) - \mathbb{H}(\mu, y_2; \mu, z_2) = \frac{\cancel{H(\mu, y_2)} - \cancel{H(\mu, y_2)} + H(\mu, x_2) - H(\mu, z_2)}{2}$$

With the same argument, we establish that there exists a function $c_2(\mu)$ such that:

$$\mathbb{H}(\mu, x_2; \mu, y_2) = \frac{H(\mu, x_2) + H(\mu, y_2)}{2} + c_2(\mu)$$

Now if we take $x_1 = y_1 = \mu$ in the first conclusion and $x_2 = y_2 = \lambda$ in the second one, we find that $c_1(\lambda) = c_2(\mu)$, $\forall\ (\lambda, \mu) \in \mathbb{R}^2$, so $c_1(\lambda) = c_2(\mu) = c \in \mathbb{R}$.

Now we derive (135) along x_2:

$$\frac{\partial \mathbb{H}}{\partial x_2}(x_1, x_2; y_1, y_2) = \frac{1}{2}\frac{\partial H}{\partial x_2}(y_1, x_2)$$

We introduce $\lambda \in \mathbb{R}$ and we integrate the previous expression between λ and x_2:

$$\mathbb{H}(x_1, x_2; y_1, y_2) - \mathbb{H}(x_1, \lambda; y_1, y_2) = \frac{1}{2}\left[H(y_1, x_2) - H(y_1, \lambda)\right]$$

We now choose $\lambda = y_2$[4]:

$$\begin{aligned}\mathbb{H}(x_1, x_2; y_1, y_2) &= \underbrace{\mathbb{H}(x_1, y_2; y_1, y_2)} + \frac{1}{2}\left[H(y_1, x_2) - H(y_1, y_2)\right]\\ &= \frac{1}{2}\left[H(x_1, y_2) - H(y_1, y_2)\right] + c \quad + \frac{1}{2}\left[H(y_1, x_2) - H(y_1, y_2)\right]\end{aligned}$$

using (136). We can simplify, and we have shown that is a function $\mathbb{H}$ satisfies (135), it must be under the form:

$$\boxed{\mathbb{H}(x_1, x_2; y_1, y_2) = \frac{1}{2}\left[H(x_1, y_2) + H(y_1, x_2)\right] + c} \tag{137}$$

Now we must verify that these functions indeed satisfy (135):

$$\begin{aligned}\mathbb{H}(x_1, x_2; y_1, y_2) - \mathbb{H}(y_1, y_2; z_1, z_2) &= \left[\frac{H(x_1, y_2) + H(y_1, x_2)}{2} + c\right] - \left[\frac{H(y_1, z_2) + H(z_1, y_2)}{2} + c\right]\\ &= \frac{H(x_1, y_2) - H(z_1, y_2) + H(y_1, x_2) - H(y_1, z_2)}{2}\end{aligned}$$

which is the expected conclusion.

The final conclusion of this paragraph is that, if $N = 2$, for any function H you can find functions $\mathbb{H}$ such that (135) is satisfied. These functions $\mathbb{H}$ that satisfy (135) are written as (137).

[4] An astute way of finding the first line without derivating is to consider $\mathbb{H}(x_1, x_2; y_1, y_2) - \mathbb{H}(x_1, y_2; y_1, y_2)$ as $[\mathbb{H}(x_1, x_2; y_1, y_2) - \mathbb{H}(y_1, y_2; x_1, y_2)] + [\mathbb{H}(y_1, y_2; x_1, y_2) - \mathbb{H}(x_1, y_2; y_1, y_2)]$ and use (135) on each bracket.

A.3 Naïve scheme for $N = 3$: a restrictive condition on H

The problem (131) becomes : Find a function $\mathbb{H} : \mathbb{R}^6 \to \mathbb{R}$ such that, $\forall\, (x_1, x_2, x_3, y_1, y_2, y_3, z_1, z_2, z_3) \in \mathbb{R}^9$,

$$\begin{aligned}\mathbb{H}(x_1, x_2, x_3; y_1, y_2, y_3) - \mathbb{H}(y_1, y_2, y_3; z_1, z_2, z_3) = &\frac{H(x_1, y_2, y_3) - H(z_1, y_2, y_3)}{2} \\ &+ \frac{H(y_1, x_2, y_3) - H(y_1, z_2, y_3)}{2} \\ &+ \frac{H(y_1, y_2, x_3) - H(y_1, y_2, z_3)}{2}\end{aligned} \tag{138}$$

Since this constraint must be true for all values of its arguments, we can first take $x_3 = y_3 = z_3 = \lambda$, and we obtain:

$$\begin{aligned}\mathbb{H}(x_1, x_2, \lambda; y_1, y_2, \lambda) - \mathbb{H}(y_1, y_2, \lambda; z_1, z_2, \lambda) = &\frac{H(x_1, y_2, \lambda) - H(z_1, y_2, \lambda)}{2} \\ &+ \frac{H(y_1, x_2, \lambda) - H(y_1, z_2, \lambda)}{2} \\ &+ \frac{\cancel{H(y_1, y_2, \lambda)} - \cancel{H(y_1, y_2, \lambda)}}{2}\end{aligned}$$

Some of the terms disappear, and we have to deal with a problem similar to (135). We can use the conclusion of the previous paragraph to establish that there exists a function $c_1(\lambda)$ such that:

$$\mathbb{H}(x_1, x_2, \lambda; y_1, y_2, \lambda) = \frac{1}{2}\left[H(x_1, y_2, \lambda) + H(y_1, x_2, \lambda)\right] + c_1(\lambda) \tag{139}$$

Using the same argument with $x_2 = y_2 = z_2 = \mu$ and $x_1 = y_1 = z_1 = \nu$, we obtain that there exist two functions $c_2(\mu)$ and $c_3(\nu)$ such that:

$$\mathbb{H}(x_1, \mu, x_3; y_1, \mu, y_3) = \frac{1}{2}\left[H(x_1, \mu, y_3) + H(y_1, \mu, x_3)\right] + c_2(\mu) \tag{140}$$

and

$$\mathbb{H}(\nu, x_2, x_3; \nu, y_2, y_3) = \frac{1}{2}\left[H(\nu, x_2, y_3) + H(\nu, y_2, x_3)\right] + c_3(\nu) \tag{141}$$

If we take $x_3 = y_3 = z_3 = \lambda$, $x_2 = y_2 = z_2 = \mu$ and $x_1 = y_1 = z_1 = \nu$ in (139), (140) and (141), we obtain that $c_1(\lambda) = c_2(\mu) = c_3(\nu)$, $\forall\, (\lambda, \mu, \nu) \in \mathbb{R}^3$, so $c_1(\lambda) = c_2(\mu) = c_3(\nu) = c \in \mathbb{R}$. Now we derive (138) along x_1:

$$\frac{\partial \mathbb{H}}{\partial x_1}(x_1, x_2, x_3; y_1, y_2, y_3) = \frac{1}{2}\frac{\partial H}{\partial x_1}(x_1, y_2, y_3) \tag{142}$$

We introduce $\nu \in \mathbb{R}$ and integrate the previous expression between ν and x_1:

$$\mathbb{H}(x_1, x_2, x_3; y_1, y_2, y_3) = \mathbb{H}(\nu, x_2, x_3; y_1, y_2, y_3) + \frac{1}{2}\left[H(x_1, y_2, y_3) - H(\nu, y_2, y_3)\right]$$

We now choose $\nu = y_1$[5]:

$$\begin{aligned}\mathbb{H}(x_1, x_2, x_3; y_1, y_2, y_3) &= \underbrace{\mathbb{H}(y_1, x_2, x_3; y_1, y_2, y_3)} + \frac{1}{2}\left[H(x_1, y_2, y_3) - H(y_1, y_2, y_3)\right] \\ &= \frac{1}{2}\left[H(y_1, x_2, y_3) + H(y_1, y_2, x_3)\right] + c + \frac{1}{2}\left[H(x_1, y_2, y_3) - H(y_1, y_2, y_3)\right]\end{aligned}$$

[5] An astute way of finding the first line without derivating is to consider $\mathbb{H}(x_1, x_2, x_3; y_1, y_2, y_3) - \mathbb{H}(y_1, x_2, x_3; y_1, y_2, y_3)$ as $[\mathbb{H}(x_1, x_2, x_3; y_1, y_2, y_3) - \mathbb{H}(y_1, y_2, y_3; y_1, x_2, x_3)] + [\mathbb{H}(y_1, y_2, y_3; y_1, x_2, x_3) - \mathbb{H}(y_1, x_2, x_3; y_1, y_2, y_3)]$ and use (138) on each bracket.

using (141). We can now write that if a function $\mathbb{H}$ satisfies (138) then it is under the form:

$$\mathbb{H}(x_1,x_2,x_3;y_1,y_2,y_3) = \frac{1}{2}\left[H(y_1,x_2,y_3)+H(y_1,y_2,x_3)+H(x_1,y_2,y_3)-H(y_1,y_2,y_3)\right]+c \tag{143}$$

Now we must verify that this kind of functions satisfies indeed (138). We have on the one hand:

$$\begin{aligned}\mathbb{H}(x_1,x_2,x_3;y_1,y_2,y_3)-\mathbb{H}(y_1,y_2,y_3;z_1,z_2,z_3) = {} & \frac{H(y_1,x_2,y_3)+H(y_1,y_2,x_3)}{2} \\ & + \frac{H(x_1,y_2,y_3)-H(y_1,y_2,y_3)}{2}+\cancel{c} \\ & - \Big(\frac{H(z_1,y_2,z_3)+H(z_1,z_2,y_3)}{2} \\ & \quad + \frac{H(y_1,z_2,z_3)-H(z_1,z_2,z_3)}{2}+\cancel{c}\Big)\end{aligned} \tag{144}$$

which must be equal to:

$$\frac{H(x_1,y_2,y_3)-H(z_1,y_2,y_3)}{2}+\frac{H(y_1,x_2,y_3)-H(y_1,z_2,y_3)}{2}+\frac{H(y_1,y_2,x_3)-H(y_1,y_2,z_3)}{2}$$

A necessary condition on H so that there exists a function $\mathbb{H}$ that satisfies (138) is that:

$$\begin{aligned}-H(y_1,y_2,y_3)-H(z_1,y_2,z_3)-H(z_1,z_2,y_3)-H(y_1,z_2,z_3)+H(z_1,z_2,z_3) = {} \\ -H(z_1,y_2,y_3)-H(y_1,z_2,y_3)-H(y_1,y_2,z_3)\end{aligned}$$

Or:

$$\begin{aligned}H(y_1,y_2,y_3)+H(z_1,y_2,z_3)+H(z_1,z_2,y_3)+H(y_1,z_2,z_3) = {} \\ H(z_1,z_2,z_3)+H(z_1,y_2,y_3)+H(y_1,z_2,y_3)+H(y_1,y_2,z_3)\end{aligned} \tag{145}$$

Since this condition must be true for any $(y_1,y_2,y_3;z_1,z_2,z3)\in\mathbb{R}^6$, it is true in particular for $z_1=y_1+h_1,\ z_2=y_2+h_2,\ z_3=y_3+h_3$ with $(h_1,h_2,h_3)\in\mathbb{R}^3$:

$$\begin{aligned}H(y_1,y_2,y_3)+H(y_1+h_1,y_2,y_3+h_3)+H(y_1+h_1,y_2+h_2,y_3)+H(y_1,y_2+h_2,y_3+h_3) = {} \\ H(y_1+h_1,y_2+h_2,y_3+h_3)+H(y_1+h_1,y_2,y_3)+H(y_1,y_2+h_2,y_3)+H(y_1,y_2,y_3+h_3)\end{aligned}$$

We group together the terms with the same color, divide by h_1:

$$\begin{aligned}&\frac{H(y_1,y_2,y_3)-H(y_1+h_1,y_2,y_3)}{h_1}+\frac{H(y_1+h_1,y_2,y_3+h_3)-H(y_1,y_2,y_3+h_3)}{h_1}+ \\ &\frac{H(y_1+h_1,y_2+h_2,y_3)-H(y_1,y_2+h_2,y_3)}{h_1}+ \\ &\frac{H(y_1,y_2+h_2,y_3+h_3)-H(y_1+h_1,y_2+h_2,y_3+h_3)}{h_1}=0\end{aligned}$$

Now we take the limit when $h_1\to 0$:

$$-\partial_1 H(y_1,y_2,y_3)+\partial_1 H(y_1,y_2,y_3+h_3)+\partial_1 H(y_1,y_2+h_2,y_3)-\partial_1 H(y_1,y_2+h_2,y_3+h_3)=0$$

We divide by h_2:

$$\frac{\partial_1 H(y_1,y_2+h_2,y_3)-\partial_1 H(y_1,y_2,y_3)}{h_2}+\frac{\partial_1 H(y_1,y_2,y_3+h_3)-\partial_1 H(y_1,y_2+h_2,y_3+h_3)}{h_2}=0$$

Then take the limit when $h_2\to 0$:

$$\partial^2_{2,1}H(y_1,y_2,y_3)-\partial^2_{2,1}H(y_1,y_2,y_3+h_3)=0$$

We divide by h_3:

$$\frac{\partial^2_{2,1}H(y_1,y_2,y_3+h_3)-\partial^2_{2,1}H(y_1,y_2,y_3)}{h_3}=0$$

And take the limit when $h_3 \to 0$[6]:

$$\boxed{\partial^3_{3,2,1}H(y_1,y_2,y_3)=0} \tag{146}$$

Now we can integrate over y_3, there exists a function $\phi(y_1,y_2)$ such that:

$$\partial^2_{2,1}H(y_1,y_2,y_3)=\phi(y_1,y_2)$$

We now integrate over y_2, there exists a function $\psi(y_1,y_3)$ such that:

$$\partial_1 H(y_1,y_2,y_3)=\underbrace{\int_0^{y_2}\phi(y_1,s)ds}_{\tilde{\phi}(y_1,y_2)}+\psi(y_1,y_3)$$

We finally integrate over y_1, there exists a function $\xi(y_2,y_3)$ such that:

$$H(y_1,y_2,y_3)=\int_0^{y_1}\tilde{\phi}(s,y_2)ds+\int_0^{y_1}\psi(s,y_3)ds+\xi(y_2,y_3)$$

To summarize, a necessary condition so that (138) can be satisfied is that there exist three functions $\phi^{1,2}$, $\phi^{1,3}$ and $\phi^{2,3}$ such that:

$$\boxed{H(y_1,y_2,y_3)=\phi^{1,2}(y_1,y_2)+\phi^{2,3}(y_2,y_3)+\phi^{1,3}(y_1,y_3)} \tag{147}$$

We can verify that condition (145) is satisfied:

$$\begin{aligned}
&\cancel{\phi^{1,2}(x_1,y_2)}+\cancel{\phi^{1,3}(x_1,y_3)}+\cancel{\phi^{2,3}(y_2,y_3)}+\cancel{\phi^{1,2}(y_1,x_2)}+\cancel{\phi^{1,3}(y_1,y_3)}+\cancel{\phi^{2,3}(x_2,y_3)}\\
&+\cancel{\phi^{1,2}(y_1,y_2)}+\cancel{\phi^{1,3}(y_1,x_3)}+\cancel{\phi^{2,3}(y_2,x_3)}+\cancel{\phi^{1,2}(x_1,x_2)}+\cancel{\phi^{1,3}(x_1,x_3)}+\cancel{\phi^{2,3}(x_2,x_3)}\\
&=\cancel{\phi^{1,2}(y_1,y_2)}+\cancel{\phi^{1,3}(y_1,y_3)}+\cancel{\phi^{2,3}(y_2,y_3)}+\cancel{\phi^{1,2}(y_1,x_2)}+\cancel{\phi^{1,3}(y_1,x_3)}+\cancel{\phi^{2,3}(x_2,x_3)}\\
&+\cancel{\phi^{1,2}(x_1,y_2)}+\cancel{\phi^{1,3}(x_1,x_3)}+\cancel{\phi^{2,3}(y_2,x_3)}+\cancel{\phi^{1,2}(x_1,x_2)}+\cancel{\phi^{1,3}(x_1,y_3)}+\cancel{\phi^{2,3}(x_2,y_3)}
\end{aligned}$$

Now, with this condition (147) fulfilled, if a function $\mathbb{H}$ satisfies (138) then it is under the form:

$$\begin{aligned}
\mathbb{H}(x_1,x_2,x_3;y_1,y_2,y_3)=\frac{1}{2}\Big[&\phi^{1,2}(y_1,x_2)+\cancel{\phi^{1,3}(y_1,y_3)}+\phi^{2,3}(x_2,y_3)+\cancel{\phi^{1,2}(y_1,y_2)}\\
&+\phi^{1,3}(y_1,x_3)+\phi^{2,3}(y_2,x_3)+\phi^{1,2}(x_1,y_2)+\phi^{1,3}(x_1,y_3)\\
&+\cancel{\phi^{2,3}(y_2,y_3)}-\cancel{\phi^{1,2}(y_1,y_2)}-\cancel{\phi^{1,3}(y_1,y_3)}-\cancel{\phi^{2,3}(y_2,y_3)}\Big]+c
\end{aligned} \tag{148}$$

$$\boxed{\begin{aligned}
\mathbb{H}(x_1,x_2,x_3;y_1,y_2,y_3)=&\frac{\phi^{1,2}(y_1,x_2)+\phi^{1,2}(x_1,y_2)}{2}\\
&+\frac{\phi^{1,3}(y_1,x_3)+\phi^{1,3}(x_1,y_3)}{2}\\
&+\frac{\phi^{2,3}(x_2,y_3)+\phi^{2,3}(y_2,x_3)}{2}+c
\end{aligned}} \tag{149}$$

from (143) and (147).

[6]The same result can be achieved by derivating directly in the expression (145)

The conclusion of this paragraph is that, if $N = 3$, it is not always possible to find a function $\mathbb{H}$ such that (138) is satisfied. If the function H verifies (146), or equivalently is sum of functions of two variables, that is to say under the form (147), then the only functions $\mathbb{H}$ that satisfy (138) are written as (149), or (143) in a undeveloped form.

Remark A.1 *This condition is very restrictive on the model used. For instance, the geometrically exact model for* 3 *unknowns (planar nonlinear vibration of a string) does not satisfy it:*

$$H_{ex}(u_1, u_2, v) = \frac{1}{2}u_1^2 + \frac{1}{2}u_2^2 + \frac{1}{2}v^2 - \alpha\Big[\sqrt{u_1^2 + u_2^2 + (1+v)^2} - (1+v)\Big]$$

So

$$\partial^3_{3,2,1} H_{ex}(u_1, u_2, v) = -3\,\alpha\, u_1\, u_2\,(1+v)\Big(u_1^2 + u_2^2 + (1+v)^2\Big)^{-3/2}$$

which is not zero as soon as $\alpha \neq 0$.

A.4 Naïve scheme for $N > 3$: a restrictive condition on H

Hypothesis A.1.1 *There exists a function* $\mathbb{H} : \mathbb{R}^{2p} \to \mathbb{R}$ *such that*

$$\begin{aligned}
&\forall\ (x_k)_{1\leq k\leq p},\ (y_k)_{1\leq k\leq p},\ (z_k)_{1\leq k\leq p} \in \mathbb{R}^{3p}, \\
&\mathbb{H}\big((x_k)_{1\leq k\leq p};\ (y_k)_{1\leq k\leq p}\big) - \mathbb{H}\big((y_k)_{1\leq k\leq p};\ (z_k)_{1\leq k\leq p}\big) = \\
&\qquad\qquad \frac{1}{2}\sum_{k=1}^{p}\Big[H(x_k, (y_l)_{l\neq k}) - H(z_k, (y_l)_{l\neq k})\Big]
\end{aligned} \tag{150}$$

A.4.1 Necessary form of $\mathbb{H}$

Hypothesis A.1.2 *If hypothesis A.1.1 is true, then* $\mathbb{H}$ *can be written, with* $c \in \mathbb{R}$*:*

$$\mathbb{H}\big((x_k)_{1\leq k\leq p};\ (y_k)_{1\leq k\leq p}\big) = \frac{1}{2}\Big[\sum_{k=1}^{p} H(x_k, (y_l)_{l\neq k}) - (p-2)H\big((y_k)_{1\leq k\leq p}\big)\Big] + c \tag{151}$$

We are going to show that hypothesis A.1.2 is true for $N \geq 3$ with a proof by mathematical induction.

Previous paragraph shows that hypothesis A.1.2 is true for $N = 3$, which begins the induction.

We suppose now that hypothesis A.1.2 is true for a certain $N \geq 3$. We suppose hypothesis A.1.1 for $p = N + 1$, i.e. there exists a function $\mathbb{H}$ such that (150). Then this equation is also true for $x_{N+1} = y_{N+1} = z_{N+1} = \lambda$:

$$\begin{aligned}
&\mathbb{H}\big((x_k)_{1\leq k\leq N}, \lambda;\ (y_k)_{1\leq k\leq N}, \lambda\big) - \mathbb{H}\big((y_k)_{1\leq k\leq N}, \lambda;\ (z_k)_{1\leq k\leq N}, \lambda\big) = \\
&\quad \sum_{k=1}^{N}\Big[H(x_k, (y_l)_{l\neq k}) - H(z_k, (y_l)_{l\neq k})\Big] + \Big[\cancel{H(\lambda, (y_l)_{l\neq k})} - \cancel{H(\lambda, (y_l)_{l\neq k})}\Big]
\end{aligned} \tag{152}$$

Then hypothesis A.1.1 is true for $p = N$. We can use the hypothesis A.1.2 to affirm that:

$$\mathbb{H}\big((x_k)_{1\leq k\leq N}, \lambda;\ (y_k)_{1\leq k\leq N}, \lambda\big) = \frac{1}{2}\Big[\sum_{k=1}^{N} H(x_k, (y_l)_{l\neq k}, \lambda) - (N-2)H\big((y_k)_{1\leq k\leq N}, \lambda\big)\Big] + c_{N+1}(\lambda)$$

Using this same argument for $x_i = y_i = z_i = \lambda_i$ for $i = 1, \ldots, N$, and finding for each i a function $c_i(\lambda_i)$, then taking in every line all $x_i = y_i = z_i = \lambda_i$, we can show that all functions c_i are constant. Then,

$$\mathbb{H}\big((x_k)_{1\le k\le N},\lambda;\,(y_k)_{1\le k\le N},\lambda\big)=\frac{1}{2}\Big[\sum_{k=1}^{N}H(x_k,(y_l)_{l\neq k},\lambda)-(N-2)H\big((y_k)_{1\le k\le N},\lambda\big)\Big]+c \tag{153}$$

We now derivate (150) along x_{N+1}:

$$\frac{\partial\mathbb{H}}{\partial x_{N+1}}\big((x_k)_{1\le k\le N+1};(y_k)_{1\le k\le N+1}\big)=\frac{1}{2}\frac{\partial H}{\partial x_{N+1}}\big(x_{N+1},(y_k)_{k\neq N+1}\big)$$

We integrate it between λ and x_{N+1}:

$$\begin{aligned}\mathbb{H}\big((x_k)_{1\le k\le N+1};(y_k)_{1\le k\le N+1}\big)-\mathbb{H}\big((x_k)_{1\le k\le N},\lambda;(y_k)_{1\le k\le N+1}\big)=&\\ \frac{1}{2}\Big[H\big(x_{N+1},(y_k)_{k\neq N+1}\big)-H\big(\lambda,(y_k)_{k\neq N+1}\big)\Big]&\end{aligned}$$

If we take $\lambda=y_{N+1}$ we can use (153):

$$\begin{aligned}\mathbb{H}\big((x_k)_{1\le k\le N+1};(y_k)_{1\le k\le N+1}\big)&=\mathbb{H}\big((x_k)_{1\le k\le N},y_{N+1};(y_k)_{1\le k\le N+1}\big)\\&\qquad+\frac{1}{2}\Big[H\big(x_{N+1},(y_k)_{k\neq N+1}\big)-H\big(y_{N+1},(y_k)_{k\neq N+1}\big)\Big]\\&=\frac{1}{2}\Big[\sum_{k=1}^{N}H(x_k,(y_l)_{l\neq k},y_{N+1})-(N-2)H\big((y_k)_{1\le k\le N},y_{N+1}\big)\Big]+c\\&\qquad+\frac{1}{2}\Big[H\big(x_{N+1},(y_k)_{k\neq N+1}\big)-H\big(y_{N+1},(y_k)_{k\neq N+1}\big)\Big]\\&=\frac{1}{2}\Big[\sum_{k=1}^{N+1}H(x_k,(y_l)_{l\neq k})-(N-1)H\big((y_k)_{1\le k\le N+1}\big)\Big]+c\end{aligned}$$

Then the hypothesis A.1.2 is true for $p=N+1$. It is then true for any $N\ge 3$.

A.4.2 Condition on H of the existence of $\mathbb{H}$

If $\mathbb{H}$ exists, we have seen in the previous paragraph that it must be under the form (151). Now we can see under which condition on H this expression satisfies indeed (150).

$$\begin{aligned}&\frac{1}{2}\Big[\cancel{\sum_{k=1}^{N}H(x_k,(y_l)_{l\neq k})}-(N-2)H\big((y_k)_{1\le k\le N}\big)\Big]+\cancel{c}\\&\qquad-\frac{1}{2}\Big[\sum_{k=1}^{N}H(y_k,(z_l)_{l\neq k})-(N-2)H\big((z_k)_{1\le k\le N}\big)\Big]-\cancel{c}\\&\qquad\qquad=\frac{1}{2}\Big[\cancel{\sum_{k=1}^{N}H(x_k,(y_l)_{l\neq k})}-\sum_{k=1}^{N}H(z_k,(y_l)_{l\neq k})\Big]\end{aligned}\tag{154}$$

Which leads to

$$(N-2)H\big((y_k)_{1\le k\le N}\big)+\sum_{k=1}^{N}H(y_k,(z_l)_{l\neq k})-(N-2)H\big((z_k)_{1\le k\le N}\big)=\sum_{k=1}^{N}H(z_k,(y_l)_{l\neq k})$$

We now choose three distinct indices k_1^*, k_2^* an k_2^*, and call $\{k_1^*,k_2^*,k_3^*\}=\mathcal{I}_*$. We call the set of all the other indices $\mathcal{I}_*^c$. Of course, $\mathcal{I}_*+\mathcal{I}_*^c=[\,1,\ldots,N]$. In the previous expression, we take

$x_k = y_k = z_k = \lambda_k$ for all $k \in \mathcal{I}_*^c$:

$$(N-2)H\big((y_k)_{k\in\mathcal{I}_*},(y_l)_{l\in\mathcal{I}_*^c}\big) + \sum_{k\in\mathcal{I}_*} H(y_k,(z_l)_{l\neq k}) + \sum_{k\in\mathcal{I}_*^c} H(y_k,(z_l)_{l\neq k}) - (N-2)H\big((z_k)_{k\in\mathcal{I}_*},(z_k)_{k\in\mathcal{I}_*^c}\big) = \sum_{k\in\mathcal{I}_*} H(z_k,(y_l)_{l\neq k}) + \sum_{k\in\mathcal{I}_*^c} H(z_k,(y_l)_{l\neq k}) \tag{155}$$

$$(N-2)H\big(y_{k_1^*},y_{k_2^*},y_{k_3^*},(\overset{\lambda_l}{\cancel{z_l}})_{l\in\mathcal{I}_*^c}\big) + H(y_{k_1^*},z_{k_2^*},z_{k_3^*},(\overset{\lambda_l}{\cancel{z_l}})_{l\in\mathcal{I}_*^c}) + H(z_{k_1^*},y_{k_2^*},z_{k_3^*},(\overset{\lambda_l}{\cancel{z_l}})_{l\in\mathcal{I}_*^c}) + H(z_{k_1^*},z_{k_2^*},y_{k_3^*},(\overset{\lambda_l}{\cancel{z_l}})_{l\in\mathcal{I}_*^c}) + \sum_{k\in\mathcal{I}_*^c} H(z_{k_1^*},z_{k_2^*},z_{k_3^*},\overset{\lambda_k}{\cancel{y_k}},(\overset{\lambda_l}{\cancel{z_l}})_{\substack{l\in\mathcal{I}_*^c\\ l\neq k}}) - (N-2)H\big(z_{k_1^*},z_{k_2^*},z_{k_3^*},(\overset{\lambda_l}{\cancel{z_l}})_{l\in\mathcal{I}_*^c}\big) = H(z_{k_1^*},y_{k_2^*},y_{k_3^*},(\overset{\lambda_l}{\cancel{y_l}})_{l\neq k}) + H(y_{k_1^*},z_{k_2^*},y_{k_3^*},(\overset{\lambda_l}{\cancel{y_l}})_{l\neq k}) + H(y_{k_1^*},y_{k_2^*},z_{k_3^*},(\overset{\lambda_l}{\cancel{y_l}})_{l\neq k}) + \sum_{k\in\mathcal{I}_*^c} H(y_{k_1^*},y_{k_2^*},y_{k_3^*},\overset{\lambda_k}{\cancel{z_k}},(\overset{\lambda_l}{\cancel{z_l}})_{\substack{l\in\mathcal{I}_*^c\\ l\neq k}}) \tag{156}$$

The first red term is multiplied by $(N-2)$, the second red term is the same term summed $N-3$ times. There is only one term left. The same thing happens with blue terms. We have:

$$H\big(y_{k_1^*},y_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}\big) + H(y_{k_1^*},z_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) + H(z_{k_1^*},y_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) + H(z_{k_1^*},z_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - H(z_{k_1^*},z_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) = H(z_{k_1^*},y_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) + H(y_{k_1^*},z_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) + H(y_{k_1^*},y_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) \tag{157}$$

We take now $z_k = y_k + h_k$ for all $k \in \mathcal{I}_*$, group similar terms and divide by $h_{k_1^*}$:

$$-\frac{H(z_{k_1^*},y_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - H\big(y_{k_1^*},y_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}\big)}{h_{k_1^*}} - \frac{H(z_{k_1^*},z_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - H(y_{k_1^*},z_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c})}{h_{k_1^*}} + \frac{H(z_{k_1^*},y_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - H(y_{k_1^*},y_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c})}{h_{k_1^*}} + \frac{H(z_{k_1^*},z_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - H(y_{k_1^*},z_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c})}{h_{k_1^*}} = 0 \tag{158}$$

We make $h_{k_1^*} \to 0$, and we get:

$$-\,\partial_{k_1^*}H(y_{k_1^*},y_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - \partial_{k_1^*}H(y_{k_1^*},z_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) + \partial_{k_1^*}H(y_{k_1^*},y_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) + \partial_{k_1^*}H(y_{k_1^*},z_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) = 0$$

We now divide by $h_{k_2^*}$ and group similar terms together:

$$\frac{\partial_{k_1^*}H(y_{k_1^*},z_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - \partial_{k_1^*}H(y_{k_1^*},y_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c})}{h_{k_2^*}} - \frac{\partial_{k_1^*}H(y_{k_1^*},z_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - \partial_{k_1^*}H(y_{k_1^*},y_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c})}{h_{k_2^*}} = 0$$

We make $h_{k_2^*} \to 0$, and we get:

$$\partial^2_{k_2^*,k_1^*}H(y_{k_1^*},y_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) - \partial^2_{k_2^*,k_1^*}H(y_{k_1^*},y_{k_2^*},z_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) = 0$$

We finally divide by $h_{k_3^*}$, take the limit when $h_{k_3^*} \to 0$ and we get:

$$\boxed{\partial^3_{k_3^*,k_2^*,k_1^*}H(y_{k_1^*},y_{k_2^*},y_{k_3^*},(\lambda_l)_{l\in\mathcal{I}_*^c}) = 0} \tag{159}$$

Since it is true for any distinct k_1^*, k_2^*, k_3^*, it means that any third crossed derivative must vanish, i.e. that H is a sum of functions of two variables:

$$\boxed{H\big((x_k)_{1\leq k\leq N}\big) = \sum_{k_1<k_2} \phi^{k_1,k_2}(x_{k_1}, x_{k_2})} \tag{160}$$

We can consequently simplify the expression of $\mathbb{H}$:

$$\begin{aligned}
\mathbb{H}\big((x_k)_{1\leq k\leq N};(y_k)_{1\leq k\leq N}\big) &= \frac{1}{2}\Big[\sum_{k=1}^{N} H\big(x_k,(y_l)_{l\neq k}\big) - (N-2)H\big((y_k)_k\big)\Big] + c \\
&= \frac{1}{2}\Big[\sum_{k=1}^{N}\Big[\sum_{\substack{k_1<k_2\\k_1=k}} \phi^{k_1,k_2}(x_{k_1}, y_{k_2}) + \sum_{\substack{k_1<k_2\\k_2=k}} \phi^{k_1,k_2}(y_{k_1}, x_{k_2}) \\
&\qquad + \sum_{\substack{k_1<k_2\\k_1\neq k\\k_2\neq k}} \phi^{k_1,k_2}(y_{k_1}, y_{k_2})\Big] - (N-2)\sum_{k_1<k_2} \phi^{k_1,k_2}(y_{k_1}, y_{k_2})\Big] + c
\end{aligned}$$

We introduce the function $\alpha(k_1, k_2, k)$:

$$\alpha(k_1,k_2,k) = \begin{cases} 0 & \text{if } k_1 = k \\ 0 & \text{if } k_2 = k \\ 1 & \text{if } k_1 \neq k \text{ and } k_2 \neq k \end{cases}$$

Which leads to the following expression for $\mathbb{H}$, where $\delta_{a,b}$ stands for the Kronecker symbol:

$$\begin{aligned}
\mathbb{H}\big((x_k)_k;(y_k)_k\big) &= \frac{1}{2}\Big[\sum_{k=1}^{N}\Big[\sum_{k_1<k_2} \delta_{k_1,k}\ \phi^{k_1,k_2}(x_{k_1}, y_{k_2}) + \sum_{k_1<k_2} \delta_{k_2,k}\ \phi^{k_1,k_2}(y_{k_1}, x_{k_2})\Big] \\
&\qquad + \sum_{k=1}^{N}\sum_{k_1<k_2} \alpha(k_1,k_2,k)\phi^{k_1,k_2}(y_{k_1}, y_{k_2}) - (N-2)\sum_{k_1<k_2} \phi^{k_1,k_2}(y_{k_1}, y_{k_2})\Big] + c \\
&= \frac{1}{2}\Big[\sum_{k_1<k_2}\Big[\phi^{k_1,k_2}(x_{k_1}, y_{k_2})\Big(\underbrace{\sum_{k=1}^{N}\delta_{k_1,k}}_{1}\Big) + \phi^{k_1,k_2}(y_{k_1}, x_{k_2})\Big(\underbrace{\sum_{k=1}^{N}\delta_{k_2,k}}_{1}\Big)\Big] \\
&\qquad \sum_{k_1<k_2} \phi^{k_1,k_2}(y_{k_1}, y_{k_2}) \underbrace{\sum_{k=1}^{N}\alpha(k_1,k_2,k)}_{\substack{N-2\\ \text{since } \alpha \text{ is 1 whenever } k\\ \text{is different from } k_1 \text{ or } k_2.}} -(N-2)\sum_{k_1<k_2} \phi^{k_1,k_2}(y_{k_1}, y_{k_2})\Big] + c
\end{aligned}$$

$$\boxed{\mathbb{H}\big((x_k)_k;(y_k)_k\big) = \sum_{k_1<k_2} \frac{\phi^{k_1,k_2}(x_{k_1}, y_{k_2}) + \phi^{k_1,k_2}(y_{k_1}, x_{k_2})}{2} + c} \tag{161}$$

We can finally verify that this expression satisfies (150). The left hand side is:

$$\begin{aligned}
\mathbb{H}\big((x_k),(y_k)\big) - \mathbb{H}\big((y_k),(z_k)\big) &= \sum_{k_1<k_2} \frac{\phi^{k_1,k_2}(x_{k_1}, y_{k_2}) + \phi^{k_1,k_2}(y_{k_1}, x_{k_2})}{2} + \cancel{c} \\
&\quad - \sum_{k_1<k_2} \frac{\phi^{k_1,k_2}(y_{k_1}, z_{k_2}) + \phi^{k_1,k_2}(z_{k_1}, y_{k_2})}{2} - \cancel{c}
\end{aligned} \tag{162}$$

And the right hand side is:

$$\frac{1}{2}\sum_{k=1}^{N}\Big[H\big(x_k,(y_l)_{l\neq k}\big)-H\big(w_k,(y_l)_{l\neq k}\big)\Big]=$$

$$=\frac{1}{2}\sum_{k=1}^{N}\Big[\big[\sum_{\substack{k_1<k_2\\k_1=k}}\phi^{k_1,k_2}(x_{k_1},y_{k_2})+\sum_{\substack{k_1<k_2\\k_2=k}}\phi^{k_1,k_2}(y_{k_1},x_{k_2})+\cancel{\sum_{\substack{k_1<k_2\\k_1\neq k\\k_2\neq k}}\phi^{k_1,k_2}(y_{k_1},y_{k_2})}\big]$$

$$-\big[\sum_{\substack{k_1<k_2\\k_1=k}}\phi^{k_1,k_2}(z_{k_1},y_{k_2})+\sum_{\substack{k_1<k_2\\k_2=k}}\phi^{k_1,k_2}(y_{k_1},z_{k_2})+\cancel{\sum_{\substack{k_1<k_2\\k_1\neq k\\k_2\neq k}}\phi^{k_1,k_2}(y_{k_1},y_{k_2})}\big]\Big]$$

$$=\sum_{k_1<k_2}\frac{\phi^{k_1,k_2}(x_{k_1},y_{k_2})\overbrace{\sum_{k=1}^{N}\delta_{k_1,k}}^{1}+\phi^{k_1,k_2}(y_{k_1},x_{k_2})\overbrace{\sum_{k=1}^{N}\delta_{k_2,k}}^{1}}{2}$$

$$-\sum_{k_1<k_2}\frac{\phi^{k_1,k_2}(z_{k_1},y_{k_2})\overbrace{\sum_{k=1}^{N}\delta_{k_1,k}}^{1}+\phi^{k_1,k_2}(y_{k_1},z_{k_2})\overbrace{\sum_{k=1}^{N}\delta_{k_2,k}}^{1}}{2}$$

$$\frac{1}{2}\sum_{k=1}^{N}\Big[H\big(x_k,(y_l)_{l\neq k}\big)-H\big(w_k,(y_l)_{l\neq k}\big)\Big]=\sum_{k_1<k_2}\frac{\phi^{k_1,k_2}(x_{k_1},y_{k_2})+\phi^{k_1,k_2}(y_{k_1},x_{k_2})}{2}$$

$$-\sum_{k_1<k_2}\frac{\phi^{k_1,k_2}(y_{k_1},z_{k_2})+\phi^{k_1,k_2}(z_{k_1},y_{k_2})}{2}$$

Which coincides with (162).

The final conclusion of this appendix is that, if $N \geq 3$, it is not always possible to find a function $\mathbb{H}$ such that

$$\forall\ (x_k)_{1\leq k\leq p},\ (y_k)_{1\leq k\leq p},\ (z_k)_{1\leq k\leq p} \in \mathbb{R}^{3p}, \tag{150}$$

$$\mathbb{H}\big((x_k)_{1\leq k\leq p};\ (y_k)_{1\leq k\leq p}\big) - \mathbb{H}\big((y_k)_{1\leq k\leq p};\ (z_k)_{1\leq k\leq p}\big) = \frac{1}{2}\sum_{k=1}^{p}\Big[H(x_k,(y_l)_{l\neq k}) - H(z_k,(y_l)_{l\neq k})\Big]$$

is satisfied. If the function H verifies

$$\partial^3_{k_3^*,k_2^*,k_1^*} H(y_{k_1^*}, y_{k_2^*}, y_{k_3^*}, (\lambda_l)_{l\in\mathcal{I}_*^c}) = 0 \tag{159}$$

or equivalently is sum of functions of two variables, that is to say under the form

$$H\big((x_k)_{1\leq k\leq N}\big) = \sum_{k_1<k_2} \phi^{k_1,k_2}(x_{k_1}, x_{k_2}) \tag{160}$$

then the only functions $\mathbb{H}$ that satisfy (150) are written as

$$\boxed{\mathbb{H}\big((x_k)_k;(y_k)_k\big) = \sum_{k_1<k_2} \frac{\phi^{k_1,k_2}(x_{k_1}, y_{k_2}) + \phi^{k_1,k_2}(y_{k_1}, x_{k_2})}{2} + c} \tag{161}$$

or in a undeveloped form:

$$\mathbb{H}\big((x_k)_{1\leq k\leq p};\ (y_k)_{1\leq k\leq p}\big) = \frac{1}{2}\Big[\sum_{k=1}^{p} H(x_k,(y_l)_{l\neq k}) - (p-2)H\big((y_k)_{1\leq k\leq p}\big)\Big] + c \tag{151}$$

B Construction of the Bank & Sujbert model

We want to show that in the case of transverse solicitations of the nonlinear string, the approximate model found when neglecting small terms is indeed the Bank-Sujbert model.

Denoting $\mathbf{u}^\varepsilon = (u^\varepsilon, v^\varepsilon)$ and

$$H(u,v) = \frac{1}{2}\,u^2 + \frac{1}{2}\,v^2 - \alpha\Big[\sqrt{u^2+(1+v)^2} - (1+v)\Big],$$

We are interested in the Cauchy problem:

$$\boxed{\left\{\begin{array}{ll} \dfrac{\partial^2 \mathbf{u}^\varepsilon}{\partial t^2} - \dfrac{\partial}{\partial x}\left[\nabla H(\dfrac{\partial \mathbf{u}^\varepsilon}{\partial x})\right] = 0, & \\ u^\varepsilon(t=0,x) = \varepsilon\,\overline{u}(x)\,, & \partial_t u^\varepsilon(t=0,x) = \varepsilon\,\overline{u}_t, \\ v^\varepsilon(t=0,x) = 0\,, & \partial_t v^\varepsilon(t=0,x) = 0. \end{array}\right.} \tag{163}$$

We seek the solution $\mathbf{u}^\varepsilon$ under the form:

$$\left|\begin{array}{l} u^\varepsilon = \varepsilon\,u_1 + \varepsilon^2\,u_2 + \varepsilon^3\,u_3 + \dots \\ v^\varepsilon = \varepsilon\,v_1 + \varepsilon^2\,v_2 + \varepsilon^3\,v_3 + \dots \end{array}\right.$$

And we want to know which system satisfy u^ε and v^ε if we neglect ε^4 in the equations.

Developing H up to order 4, we obtain:

$$H_{DL4}(u,v) = \frac{1-\alpha}{2}\,u^2 + \frac{1}{2}\,v^2 + \frac{\alpha}{2}\,u^2 v + \frac{\alpha}{8}\,u^4 - \frac{\alpha}{2}\,u^2 v^2$$

Hence

$$\nabla H_{DL4}(u,v) = \begin{pmatrix} (1-\alpha)\,u + \alpha\,uv + \dfrac{\alpha}{2}\,u^3 - \alpha\,uv^2 \\ v + \dfrac{\alpha}{2}\,u^2 - \alpha\,u^2 v \end{pmatrix} \tag{164}$$

It is easy to see that

$$\left|\begin{array}{l} (u^\varepsilon)^2 = \varepsilon^2 u_1^2 + 2\varepsilon^3 u_1 u_2 + \mathcal{O}(\varepsilon^4) \\ (v^\varepsilon)^2 = \varepsilon^2 v_1^2 + 2\varepsilon^3 v_1 v_2 + \mathcal{O}(\varepsilon^4) \\ (u^\varepsilon)^2 v^\varepsilon = \varepsilon^3 u_1^2 v_1 + \mathcal{O}(\varepsilon^4) \\ u^\varepsilon v^\varepsilon = \varepsilon^2 u_1 v_1 + \varepsilon^3 (u_1 v_2 + u_2 v_1) + \mathcal{O}(\varepsilon^4) \\ u^\varepsilon (v^\varepsilon)^2 = \varepsilon^3 u_1 v_1^2 + \mathcal{O}(\varepsilon^4) \\ (u^\varepsilon)^3 = \varepsilon^3 u_1^3 + \mathcal{O}(\varepsilon^4) \end{array}\right.$$

☛ We can write the part of the Cauchy problem (163) which is in factor of ε:

$$(\mathcal{S}_1) \quad \left\{\begin{array}{ll} \dfrac{\partial^2 u_1}{\partial t^2} - (1-\alpha)\,\dfrac{\partial^2 u_1}{\partial x^2} = 0, & \\ \dfrac{\partial^2 v_1}{\partial t^2} - \dfrac{\partial^2 v_1}{\partial x^2} = 0, & \\ u_1(t=0,x) = \overline{u}(x)\,, & \partial_t u_1(t=0,x) = \overline{u}_t, \\ v_1(t=0,x) = 0\,, & \partial_t v_1(t=0,x) = 0. \end{array}\right.$$

This system decouples u_1 and v_1, and we notice that v_1 is solution of a linear wave equation with speed 1 and initial data 0 which leads to

$$\boxed{v_1 \equiv 0.}$$

On the other hand, u_1 is simply the solution of a linear wave equation with speed $1-\alpha$, and initial data $\overline{u}(x)$ and $\overline{u}_t$. The expressions are now simpler:

$$\left| \begin{array}{l} u^\varepsilon = \varepsilon\, u_1 + \varepsilon^2\, u_2 + \varepsilon^3\, u_3 + \dots \\ v^\varepsilon = \varepsilon^2\, v_2 + \varepsilon^3\, v_3 + \dots \end{array} \right.$$

and

$$\left| \begin{array}{l} (u^\varepsilon)^2 = \varepsilon^2 u_1^2 + 2\varepsilon^3 u_1 u_2 + \mathcal{O}(\varepsilon^4) \\ (v^\varepsilon)^2 = \mathcal{O}(\varepsilon^4) \\ (u^\varepsilon)^2 v^\varepsilon = \mathcal{O}(\varepsilon^4) \\ u^\varepsilon v^\varepsilon = \varepsilon^3 u_1 v_2 + \mathcal{O}(\varepsilon^4) \\ u^\varepsilon (v^\varepsilon)^2 = \mathcal{O}(\varepsilon^4) \\ (u^\varepsilon)^3 = \varepsilon^3 u_1^3 + \mathcal{O}(\varepsilon^4) \end{array} \right.$$

☛ We write the part of the Cauchy problem (163) which is in factor of ε^2:

$$(\mathcal{S}_2) \quad \left\{ \begin{array}{l} \dfrac{\partial^2 u_2}{\partial t^2} - (1-\alpha)\, \dfrac{\partial^2 u_2}{\partial x^2} = 0, \\ \dfrac{\partial^2 v_2}{\partial t^2} - \dfrac{\partial^2 v_2}{\partial x^2} + \dfrac{\alpha}{2} \dfrac{\partial}{\partial x}\Big((\dfrac{\partial u_1}{\partial x})^2 \Big) = 0, \\ u_2(t=0,x) = 0\,, \quad \partial_t u_2(t=0,x) = 0, \\ v_2(t=0,x) = 0\,, \quad \partial_t v_2(t=0,x) = 0. \end{array} \right.$$

Using the same reasoning, $u_2 \equiv 0$ and v_2 can be seen as the solution of a linear wave equation with right member $-\dfrac{\alpha}{2}\dfrac{\partial}{\partial x}\Big((\dfrac{\partial u_1}{\partial x})^2 \Big)$, which is known thanks to $(\mathcal{S}_1)$.

☛ We finally write the part of the Cauchy problem (163) which is in factor of ε^3:

$$(\mathcal{S}_3) \quad \left\{ \begin{array}{l} \dfrac{\partial^2 u_3}{\partial t^2} - (1-\alpha)\, \dfrac{\partial^2 u_3}{\partial x^2} + \alpha \dfrac{\partial}{\partial x}\Big[\dfrac{\partial u_1}{\partial x}\dfrac{\partial v_2}{\partial x} + \dfrac{1}{2}\Big(\dfrac{\partial u_1}{\partial x}\Big)^3 \Big] = 0, \\ \dfrac{\partial^2 v_3}{\partial t^2} - \dfrac{\partial^2 v_3}{\partial x^2} + \alpha \dfrac{\partial}{\partial x}\Big[\dfrac{\partial u_1}{\partial x}\dfrac{\partial u_2}{\partial x} \Big] = 0, \\ u_3(t=0,x) = 0\,, \quad \partial_t u_3(t=0,x) = 0, \\ v_3(t=0,x) = 0\,, \quad \partial_t v_3(t=0,x) = 0. \end{array} \right.$$

☛ We combine the three systems in the following way : $\varepsilon(\mathcal{S}_1) + \varepsilon^2(\mathcal{S}_2) + \varepsilon^3(\mathcal{S}_3)$:

$$(\mathcal{S}_\varepsilon) \quad \left\{ \begin{array}{l} \partial_t^2(\varepsilon u_1 + \ \varepsilon^2 u_2 + \varepsilon^3 u_3) - \partial_x \Big[(1-\alpha)\partial_x(\varepsilon u_1 + \ \varepsilon^2 u_2 + \varepsilon^3 u_3) \\ \qquad\qquad\qquad + \varepsilon^3 \alpha \partial_x u_1 \partial_x v_2 + \varepsilon^3 \dfrac{\alpha}{2}(\partial_x u_1)^3 \Big] = 0, \\ \partial_t^2(\varepsilon v_1 + \ \varepsilon^2 v_2 + \varepsilon^3 v_3) - \partial_x \Big[\partial_x(\varepsilon v_1 + \ \varepsilon^2 v_2 + \varepsilon^3 v_3) + \varepsilon^2 \dfrac{\alpha}{2}(\partial_x u_1)^2 \\ \qquad\qquad\qquad + \varepsilon^3 \alpha \partial_x u_1 \partial_x u_2 \Big] = 0, \end{array} \right.$$

But we know that

$$\left| \begin{array}{l} u^\varepsilon = \varepsilon u_1 + \varepsilon^2 u_2 + \varepsilon^3 u_3 + \mathcal{O}(\varepsilon^4) \\ v^\varepsilon = \varepsilon v_1 + \varepsilon^2 v_2 + \varepsilon^3 v_3 + \mathcal{O}(\varepsilon^4) \\ u^\varepsilon v^\varepsilon = \varepsilon^3 u_1 v_2 + \mathcal{O}(\varepsilon^4) \\ (u^\varepsilon)^3 = \varepsilon^3 u_1^3 + \mathcal{O}(\varepsilon^4) \\ (u^\varepsilon)^2 = \varepsilon^2 u_1^2 + 2\varepsilon^3 u_1 u_2 + \mathcal{O}(\varepsilon^4) \end{array} \right.$$

Which leads to the following system for $(u^\varepsilon, v^\varepsilon)$, if we neglect ε^4:

$$(\mathcal{S}_{BS}) \quad \left\{ \begin{array}{ll} \dfrac{\partial^2 u^\varepsilon}{\partial t^2} - \dfrac{\partial}{\partial x}\Big[(1-\alpha)\,\dfrac{\partial u^\varepsilon}{\partial x} + \alpha \dfrac{\partial u^\varepsilon}{\partial x}\,\dfrac{\partial v^\varepsilon}{\partial x} + \dfrac{\alpha}{2}\big(\dfrac{\partial u^\varepsilon}{\partial x}\big)^3\Big] = 0, & \\ \dfrac{\partial^2 v^\varepsilon}{\partial t^2} - \dfrac{\partial}{\partial x}\Big[\dfrac{\partial v^\varepsilon}{\partial x} + \dfrac{\alpha}{2}\big(\dfrac{\partial u^\varepsilon}{\partial x}\big)^2\Big] = 0, & \\ u^\varepsilon(t=0,x) = \varepsilon\,\overline{u}(x)\,, & \partial_t u^\varepsilon(t=0,x) = \varepsilon\,\overline{u}_t, \\ v^\varepsilon(t=0,x) = 0\,, & \partial_t v^\varepsilon(t=0,x) = 0. \end{array} \right.$$

Which corresponds to the Cauchy problem (163) with

$$H(u,v) = \frac{1-\alpha}{2},u^2 + \frac{1}{2}\,v^2 + \frac{\alpha}{2}\,u^2 v + \frac{\alpha}{8}\,u^4$$

which is exactly the empirical approximation done by Bank and Sujbert in [2].

References

[1] G V Anand. Large-amplitude damped free vibration of a stretched string. *Journal of the Acoustical Society of America*, 45(5):1089–1096, 1969.

[2] B Bank and L Sujbert. Generation of longitudinal vibrations in piano strings: From physics to sound synthesis. *Journal of the Acoustical Society of America*, 117:2268–2278, 2005.

[3] Julien Bensa. *Analyse et synthèse de sons de piano par modèles physiques et de signaux*. PhD thesis, 2003.

[4] S Bilbao. Conservative numerical methods for nonlinear strings. *Journal of the Acoustical Society of America*, 118(5):3316–3327, 2005.

[5] B Cano. Conserved quantities of some hamiltonian wave equations after full discretization. *Numerische Mathematik*, 103:197–223, 2006.

[6] Antoine Chaigne and Anders Askenfelt. Numerical simulation of piano strings. i. a physical model for a struck string usin finite-difference methods. *Acoustical Society of America Journal*, 95(2):1112–1118, Feb 1994.

[7] YS Chin and C Qint. Explicit energy-conserving schemes for the three-body problem. *Journal of Computational Physics*, 83(2):485–493, 1989.

[8] H A Conklin. Design and tone in the mechanoacoustic piano. part iii. piano strings and scale design. *Acoustical Society of America Journal*, 100:1286–1298, 1996.

[9] H A Conklin. Generation of partials due to nonlinear mixing in a stringed instrument. *Acoustical Society of America Journal*, 105:536–545, 1999.

[10] J de Frutos and JM Sanz-Serna. Accuracy and conservation properties in numerical integration: the case of the korteweg-de-vries equation. *Numerische Mathematik*, 75(4):421–445, 1997.

[11] S Dmitriev, P Kevrekidis, and N Yoshikawa. Standard neares-neighbour discretizations of klein gordon models cannot preserve both energy and linear momentum. *J Phys A-Math Gen*, 39:7217–7226, 2006.

[12] D Furihata. Finite-difference schemes for nonlinear wave equation that inherit energy conservation property. *Journal of Computational and Applied Mathematics*, 134(1-2):37–57, 2001.

[13] N Giordano and A J Korty. Motion of a piano string: Longitudinal vibrations and the role of the bridge. *Acoustical Society of America Journal*, 100(6):3899–3908, 1996.

[14] E Godlewski and P A Raviart. *Hyperbolic systems of conservation laws*. 1991.

[15] O Gonzalez. Exact energy and momentum conserving algorithms for general models in nonlinear elasticity. *Computer Methods in Applied Mechanics and Engineering*, 190:1763–1783, 2000.

[16] O Gonzalez and J Simo. On the stability of symplectic and energy-momentum algorithms for non-linear hamiltonian systems with symmetry. *Computer Methods in Applied Mechanics and Engineering*, 134:197–222, 1996.

[17] D Greenspan. Conservative numerical methods for $\ddot{x}= f(x)$. *Journal of Computational Physics*, 56:28–41, 1984.

[18] E Hairer. Backward analysis of numerical integrators and symplectic methods. *Ann. Numer. Math.*, 1(1-4):107–132, 1994.

[19] E Hairer and C Lubich. The life-span of backward error analysis for numerical integrators. *Numerische Mathematik*, 76:441–462, 1997.

[20] MW Hirsch, S Smale, and LR Devaney. Differential equations, dynamical systems, and an introduction to chaos. page 417, 2004.

[21] F John. Formation of singularities in one-dimensional nonlinear wave propagation. *Comm. Pure Appl. Math.*, 27(3):377–405, 1974.

[22] NA Kampanis, VA Dougalis, and JA Ekaterinakis. Effective computational methods for wave propagation. *Chapman & Hall / CRC*, 2008.

[23] C Kane, JE Marsden, and M Ortiz. Symplectic-energy-momentum preserving variational integrators. *Journal of Mathematical Physics*, 40:3353, 1999.

[24] P Kevrekidis. On a class of discretizations of hamiltonian nonlinear partial differential equations. *Physica D: Nonlinear Phenomena*, 183(1-2):68–86, Sep 2003.

[25] S Klainerman and A Majda. Formation of singularities for wave equations including the nonlinear vibrating string. *Comm. Pure Appl. Math.*, 33(3):241–263, 1980.

[26] E V Kurmyshev. Transverse and longitudinal mode coupling in a free vibrating soft string. *Physics Letters A*, 310(2-3):148–160, 2003.

[27] Ta-Tsien Li. *Global classical solutions for quasilinear hyperbolic systems*. 1994.

[28] RE Mickens. A non-standard finite-difference scheme for conservative oscillators. *Journal of Sound and Vibration*, 240(3):587–591, 2001.

[29] RE Mickens. A numerical integration technique for conservative oscillators combining non-standard finite-difference methods with a hamilton's principle. *Journal of Sound and Vibration*, 285(1-2):477–482, 2005.

[30] PM Morse and KU Ingard. *Theoretical Acoustics*. 1968.

[31] Isao Nakamura and D Naganuma. Characteristics of piano sound spectra. 1993.

[32] S Reich. Backward error analysis for numerical integrators. *SIAM Journal on Numerical Analysis*, 36(5):1549–1570, 1999.

[33] M B Rubin and O Gottlieb. Numerical solutions of forced vibration and whirling of a non-linear string using the theory of a cosserat point. *Journal of Sound and Vibration*, 197(1):85–101, 1996.

[34] J Sanz-Serna. Symplectic integrators for hamiltonian problems: an overview. *Acta Numerica*, 1991.

[35] W Strauss and L Vazquez. Numerical solution of a nonlinear klein-gordon equation. *Journal of Computational Physics*, 28(2):271–278, 1978.

[36] C Valette and C Cuesta. *Mécanique de la corde vibrante*. 1993.

[37] L Vu-Quoc and S Li. Invariant-conserving finite difference algorithms for the nonlinear klein-gordon equation. *Computer Methods in Applied Mechanics and Engineering*, 107(3):341–391, 1993.

[38] G Zhong and JE Marsden. Lie-poisson hamilton-jacobi theory and lie-poisson integrators. *Physics Letters A*, 133(3):134–139, Nov 1988.

Centre de recherche INRIA Paris – Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399